Running Head: USING SCENARIO-BASED TRAINING TO TEACH SRM

Using Scenario-Based Training to Teach Single Pilot Resource Management Related to the Use of the BRS Parachute

by
Shayna Strally

A thesis submitted to the

Office of Graduate Programs

in Partial Fulfillment of the Requirements for the Degree of

Master of Human Factors and Systems

Embry-Riddle Aeronautical University

Daytona Beach, FL

2005

Table of Contents

ABSTRACT

The Ballistic Recovery System is an emergency parachute for single engine aircraft which, when released, lowers the aircraft to the ground to prevent terrain collision. This study sought to examine the effects of scenario-based training on pilot's use of the BRS. Of particular interest was the point at which the pilot decides to deploy the BRS. Single pilot resource management was included as a training objective, as it encompasses relevant cognitive skills such as decision making and situational awareness. The results showed participants in the scenario-based training condition performed significantly better than participants in a traditional training condition on several measures. Although additional research is needed, these results likely indicate that scenario-based training is more effective for training emergency parachute use.

INTRODUCTION

Problem Statement

More than 20,000 general aviation accidents occurred between 1990 and 1999, with nearly 4000 of those accidents fatal according to the National Transportation Safety Board's (NTSB) aviation database. While many researchers are looking into the causes of aviation accidents and how they might be prevented, a recent innovation has provided another potential solution for the accidents that result in fatalities. Specifically, an emergency parachute for aircraft has been developed and successfully tested by the Ballistic Recovery System (BRS) company. When deployed, the parachute lowers the entire plane to the ground while maintaining an acceptable level of safety. While the probability of survival for the passengers is dramatically improved because of the parachute, the plane will sustain a significant amount of damage on impact, perhaps enough to permanently disable the aircraft. Additionally, pilots must make rapid decisions for deployment of the BRS parachute in stressful circumstances, and therefore, error is quite possible.

As the BRS parachute is new, most pilots will not have had prior training for this type of system. Hence, the development of the BRS parachute for aircraft presents the need for research to assess the impacts of training on pilots' use of the device. Indeed, the BRS is meant to be a last resort -- the only option left to the pilot for survival. The risk is that a novice or scared pilot may deploy the parachute prematurely before all reasonable options have been exhausted. On the other hand, an overconfident pilot may deploy the parachute too late. Finally, a pilot unaccustomed to the option of using a parachute may simply not think to use it at all. This may be particularly true under time pressure and stress when cognitive tunneling occurs (e.g., Cook & Woods, 1994; Orasanu, 1997). The purpose of the current research is twofold: 1) to develop and validate a scenario-based training intervention that is designed specifically for parachute deployment in the single engine aircraft, and 2) to add to the research base regarding the efficacy of scenario-based training.

This literature review will begin with a brief comparison between the BRS parachute and ejection seats. A discussion of a related research and design program will follow (FAA/Industry Training Standards, 2004). Next, the knowledge, skills, attitudes and other characteristics (KSAOs) for effective use of the BRS are discussed. Finally, the scenario-based training method and its utility for training parachute use will be discussed.

BRS Parachutes vs. Ejection Seats

Since the Ballistic Recovery System has a similar purpose to ejection seats and other safety devices of aircraft, a brief review of these areas is warranted. Ejection seats are commonly found in military aircraft. This includes both aircraft currently used by the military and some converted aircraft now used for general aviation purposes. According to Bonsor (n.d.), many different models of planes carry ejection seats because the seats are often the only alternative to save the life of a pilot faced with a damaged plane ready to crash. Ejection seats are complex devices built around typical seat components, and are activated by various means – pull handles, face curtains, etc. When activated, an explosive cartridge propels the seat through a canopy on the plane, releasing a temporary drogue parachute immediately. An altitude sensor then signals the release of the main parachute when the pilot is at a safe distance from the aircraft, lowering the pilot safely to the ground.

While the situations that require the use of the BRS parachute are similar to the situations in which pilots should eject, it is still necessary to study the specific effects of training on the BRS parachute. One reason is because the procedure to eject differs slightly from the procedure to release the parachute. In addition, pilots are also instructed to eject from much greater heights. For example, some maintain a height minimum of 10,000 ft. before ejection (Callaghan & Irwin, 2003). In contrast, the minimum release height for the BRS is only 500 ft. Although these planes are obviously flown at different heights normally, this could still affect decision time, perhaps providing the pilot using the BRS with more time to contemplate the decision and carry out the release process. Another difference between ejection seats and the BRS parachute involves the pilot population. The majority of aircraft equipped with ejection seats are operated by the military. Military pilots may differ from general aviation pilots in many aspects. For example, military pilots are thought to be more willing to take risks than are commercial pilots (Sicard, Taillemite, Jouve, & Blin, 2003). This may apply to a comparison between military and general aviation pilots as well. Military pilots generally fly much more expensive aircraft, sometimes worth millions of dollars which belong to the military -- not the pilots. Therefore, they might attempt to save the aircraft at any cost and avoid ejection as much as possible. Another difference may be level of training. Private pilots generally have the least amount of training compared to military and commercial pilots who have access to recurring training programs through their respective employers

(Hunter, 1997). Military pilots may therefore be better prepared for emergencies than the average general aviation pilot.

Despite these differences, ejection studies on training and decision making are often similar to the goals of this study and are referenced in this study where appropriate.

Training is also particularly important when the equipment is unfamiliar as is the BRS parachute to most pilots. New technology often results in new equipment and altered skill requirements. This creates a need for training job-specific knowledge and the necessary skills (Noe, 1999). For these reasons, the effect of a training intervention on the use of the BRS should be examined to complement research on ejection seats.

FAA Industry Training Standards (FITS)

The FAA Industry Training Standards (FITS) is joint research venture, involving the FAA's Center for General Aviation Research, Embry-Riddle Aeronautical University, the University of North Dakota and the general aviation industry. The main objective of this venture is to "ensure pilots learn to safely, competently and efficiently operate a technically advanced piston or light jet aircraft in the modern National Airspace System (NAS)" (FAA/Industry Training Standards, 2004). Goals of FITS also include reducing general aviation accidents by a significant amount. In other words, one goal is to reduce pilot error, as the majority of GA accidents (75%) are pilot error related (FAA/Industry Training Standards, 2004). The argument is that to achieve these goals, and to account for the technically advanced aircraft (TAA) recently introduced in general aviation, a new training style must be adopted to amend this problem. Specific training goals include enhancing higher order thinking, including aeronautical decision making (ADM), situational awareness, pattern recognition and decision making (FAA/Industry Training

Standards, 2004). Other skills included within the FITS training goals are "automation competence, planning and execution, procedural knowledge, and psychomotor skills" (FAA/Industry Training Standards, 2004).

Single Pilot Resource Management

Single pilot resource management (SRM) is also included in the goals of the FITS training program, and represents an important concept in the current study. FITS defines SRM as the "art and science of managing all resources (both on-board the aircraft and from outside sources) available to a single-pilot (prior and during flight) to ensure the successful outcome of the flight is never in doubt" (FAA/Industry Training Standards, 2004). SRM is similar to crew resource management (CRM), yet while CRM has been studied in greater depth, SRM is a relatively new area. Lauber (1984) defined CRM as "using all available resources – information, equipment, and people - to achieve safe and efficient flight operations". This is closely related to the FITS definition for SRM. The main difference is, of course, SRM reflects the activities of only one pilot flying an aircraft as is usually the case in general aviation, while CRM considers each individual but also the interactions between crewmembers. The application of CRM to the single pilot has not been studied. Due to the lack of research on SRM, available research on CRM will also be reviewed here.

Although there is some dispute regarding the specific behaviors which comprise CRM, it is generally thought to involve six main behaviors: situational awareness, crew coordination/flight integrity, communication, risk management and decision making, task management, mission planning/debrief (Karp & Nullmeyer, 2001). This definition of CRM is based on research within commercial aviation; therefore, some of the behaviors are not relevant to a single pilot aircraft (e.g. crew coordination). Communication also may not play a very large role with a single pilot, yet the pilot must communicate periodically with ATC, as well as with the passengers (if any are on board) and interact with equipment. However, other CRM behaviors - including SA, flight integrity, and risk management and decision making - are essential for any pilot. Turner (1995) asserts that with some translation, the CRM principles are applicable to single pilots in general aviation.

Wilson-Donnelly & Shappell (2004) associate somewhat different concepts with CRM, based on a review of the causal factors attributed to CRM related U.S. Navy/Marine Corp aviation accidents between 1990 and 2000. Pilots were asked to classify the CRM related issues, and researchers narrowed the behaviors down to six: failure to conduct adequate briefs, lack of communication, miscommunication, failure to monitor, failure to backup/assist, and failure to utilize resources. Although the researchers maintain the importance of skills like SA and decision making, they believe these skills are separate concepts and not to be confused with CRM. Again, certain behaviors noted here (e.g. conducting adequate briefs/planning, adequately monitoring and utilizing resources) seem relevant to single pilots, but communication issues are not as prevalent. There is still some overlap between the definitions. For example, failure to monitor may be part of situational awareness, and both models emphasize problems with communication.

The FITS training program has also provided an explanation of CRM behaviors, and has adapted it to the single pilot. According to FITS, there are six main behaviors within SRM: aeronautical decision making, automation management, task management, situational awareness, risk management, CFIT avoidance (FAA/Industry Training Standards, 2004). Due to the lack of research on SRM, and the agreed upon importance of SA and DM in aviation, the explanation of SRM adapted by FITS will be utilized in this study.

FITS also offers a technique that may be used to enhance SRM, referred to as "the 5 P's" (FAA/Industry Training Standards, 2005). The 5 P's represent the plan, plane, pilot, passengers, and programming. Pilots are encouraged to check of the status of the 5 P's at certain decision points throughout the flight, including: before leaving the flight planning room, before leaving the ground, every second fuel check, before leaving cruise altitude, and before leaving the IAF.

Failed competencies of CRM behaviors, including incomplete situational awareness, poor communication/coordination and inadequate planning, have caused many aviation accidents (Karp & Nullmeyer, 2001). Therefore, adequately training competency in these behaviors will ensure additional safety for pilots. Training design begins with a needs assessment which typically involves a task analysis. A task analysis describes the work activities involved and any knowledge, skills, abilities, and other characteristics (KSAOs) employees must have (Noe, 1999). A preliminary task analysis follows, to determine the specific KSAOs pilots must have to properly use the BRS parachute.

KSAOs Involved in Using the BRS

As noted earlier, it is likely that flying an aircraft with the BRS parachute requires some different KSAOs then those required to fly traditional aircraft. An essential step in developing training is to first conduct a task analysis to identify the knowledge, skills, abilities and other characteristics (KSAOs) necessary for optimal performance of the job or task (Aamodt, 2004). *Knowledge* is the information necessary for the performance of a task. A *skill* is a required proficiency of a learned task and *ability* is a basic capacity to perform a variety of tasks. "Other characteristics" include aspects such as personality, interest, motivation, licenses, degrees and experience. Every aviation task, from reading the altimeter to landing the plane in inclement weather, requires KSAOs. Correct use of the BRS parachute is no exception. The KSAOs required for successful use of the parachute may differ, at least slightly, from other aviation tasks.

Table 1 contains an initial list of tasks and related KSAOs for successful use of the BRS parachute.

***Table 1.* KSAOs for the BRS parachute**

Tasks for using the BRS		**Specific KSAOs**
Awareness of Problem	Knowledge	Aircraft malfunctions which might lead to problems - Engine failures; Weather disturbance; Equipment malfunctions Incorrect displays Common human errors (overconfidence, poor SA, misreading instrument)
	Skills	Situational awareness (one aspect of SRM) of the status of all critical elements in and outside of cockpit (with instruments, weather, etc.)

		Risk perception (one aspect of SRM) - notice any potential problems or risks (with instruments, weather, etc.)
	Abilities	Sufficient visual ability with both near and far vision
	Other	
Decision to deploy	Knowledge	Typical situations which require BRS; emergencies including: Engine failures, weather disturbance, equipment malfunctions
		Typical situations which do not require BRS; including: Normal instrument readings, minor problems, major problems if can still safely land plane
		When to deploy parachute - Min/Max feet (altitude) for safe deployment; Max speeds for safe deployment
	Skills	Situational awareness of the status of all critical elements in and outside of cockpit (with instruments, weather, etc.)
		Decision Making (one aspect of SRM) - generate at least one plausible solution (such as land plane if possible, use BRS as last resort), predict outcome to confirm the solution is safe and appropriate
		SRM, or overall resource management within the cockpit
	Abilities	Reaction time
	Other	Pilots license; appropriate instrument rating; training
Deployment	Knowledge	The steps for deploying the parachute
	Skills	
	Abilities	Manual and finger dexterity

		Reaction time
	Other	Willingness to pull before its too late

Certain KSAOs are repeated in the table, such as situational awareness and decision making. These are described it greater detail next, as they are complex issues and are also important to the overall success for use of the BRS. Although all KSAOs noted here are important to the use of the parachute, due to a lack of time and other constraints, decision-making/risk management and situational awareness are the only high order skills from the six SRM behaviors explored in greater depth in this study.

Situational Awareness

According to Endsley (2000), situational awareness is essentially "knowing what is going on around you." For pilots, this involves knowledge of important information, i.e., that which is necessary for safe flight. Poor situational awareness has been cited as a frequent cause in aviation accidents, particularly in the case of weather-related accidents in general aviation (Bell & Mauro, 2000).

SA is comprised of three interdependent levels. The first level is the perception of elements within the environment (i.e. cockpit). In the second level, comprehension, the perceived elements must be understood in order for the pilot to benefit from the perception. The third level involves the projection of the elements within the environment into future time, or the ability to predict future events. The perception of time is also important in SA and is part of levels two and three. To sum, SA is "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future" (Endsley, 1988).

Endsley (2000) also categorizes situational awareness and decision making as separate stages. However, they are strongly linked and the success of decision making often depends on whether a pilot had effective SA. A pilot may exhibit excellent SA, but still make a poor decision, and vice versa. Situational awareness is gleaned from a number of sources, where cues are perceived by the senses, both at consciously and subconsciously.

Working memory and attention both have an impact on SA, and have limited capacities. Pattern matching, mental models, and schemata may help pilots overcome memory limitations and thus improve SA.

Decision Making

Decision-making is a fundamental component of any aviation operation. One study estimated over 50% of fatal accidents may be attributed to poor decision making by the flight crew (Jensen, 1982). Decision errors are believed to cause more accidents than either procedural or manual errors. This has brought on a trend of incorporating decision making training into many pilot training programs (Klein, 2000). However, some researchers believe too much emphasis is placed on the importance of decision making in aviation accidents. For example, Shappell & Wiegmann (2001) analyzed controlled flight into terrain (CFIT) aviation accidents using the Human Factors Analysis and Classification System (HFACS). There results showed while decision errors contribute to a number of CFIT accidents, they are associated more frequently with *non*-fatal rather than fatal accidents. Fatal CFIT accidents instead appear to be more commonly associated with violations.

Although there is disagreement on the overall frequency of decision making related aviation accidents, this dispute most likely does not affect this study, as the focus here is training pilots to decide when to use the BRS successfully. Therefore, decision making is certainly relevant to the goals of this study. Also, decision errors still contribute to at least some aviation accidents. Training which might reduce accident rates by even a small percentage is worthy of attention.

The Information Processing Model as it Relates to Decision Making

In terms of the basic decision making process, a decision occurs when a decision maker must select one option from two or more possibilities, base the decision on available information within a certain time frame (must last longer than one second) and the best decision is not obvious or even certain, so therefore risk is involved (Wickens, Gordon & Liu, 1998). Deployment of the BRS parachute qualifies as a decision under these criteria.

Sturgeon (1988) presents a version of the information-processing model, identifying five basic steps which occur during decision making. A pilot first *acquires* information from the environment, through various senses (auditory, visual, tactile, etc. channels). The next step involves *perception*, where the brain determines defining characteristics of the information, such as quality and quantity, and integrates the information from the senses to ensure the information is accurate. Situational awareness is important in these first two steps, as the pilot must ultimately make an informed decision based on an accurate assessment of the environment. The brain continues to *process* the information until decision alternatives are devised in reaction to the impending characteristics of the environment. The alternatives could include the decision to eject or use the BRS parachute, to attempt to land, change altitude, speed, etc. The pilot selects one of these alternatives based on the probability of success and survival, using his/her *judgment*, based on individual characteristics such as training, personal characteristics, experience and so forth. This area of decision-making is most relevant to the goals of the training intervention described in this paper. The pilot will then *implement*, or carry out, his/her selected alternative.

Classic Decision Theory

In terms of how decision-making occurs, many decision-making theories exist on the cognitive processes involved. Proponents of classic decision theory believe in “Rational Choice” or the notion that an ideal process of decision-making exists for each and every decision. Mathematical formulas are used to select the best alternative, based on each choices relative ranking, or value (Wickens et al., 1998). A form of Rational Choice theory is the DECIDE model (Benner, 1975). The six steps of DECIDE essentially include: Detect a change in the environment; Estimate the effect of the change; Choose a safe outcome; Identify plausible alternate actions; Do the best action; Evaluate the effect of choice. The DECIDE model has been used commonly in aviation training, as well as within business and engineering classes in the educational setting.

Although the Rational Choice method is applicable in some settings, it is often not realistic in aviation (Klein, 2000), including the decision of whether or not to deploy the BRS. The time allocated for making the decision is usually not enough to perform a formal analysis. Furthermore, there is a great deal of uncertainty in the situation. Pilots under time pressure typically do not generate more alternatives after determining one good option (Klein, 1998; Klein, Orasanu, Calderwood, & Zsambok, 1993; Zsambok & Klein, 1997). For a normative model of decision making to work, the pilot/decision maker must determine the probability of success and the potential consequence of the decision within an acceptable amount of time (Wickens et al., 1998). With training and experience, a pilot's estimation of these variables may improve, but each pilots expected utility of an alternative would differ to some degree (Wickens et al., 1998).

Naturalistic Decision Making

Naturalistic Decision Making (NDM) theory may better describe the decision making processes occurring in aviation, particularly during emergencies (Wickens et al., 1998; Klein, 2000; Zsambok & Klein, 1997). Zsambok (1997) defines NDM as "the way people use their experience to make decisions in field settings." Experience and knowledge are important factors in NDM (Orasanu & Connolly, 1993). Associated research examines how people develop and use both factors to make decisions (Klein, 2000). NDM focuses on decisions made in the "field", or real world, which tend to occur in constantly changing environments, have poorly structured and incomplete information, changing goals, limited time, multiple decision makers and a high degree of risk (Orasanu & Connolly, 1993). These factors make decision making more difficult (Klein, 2000). It is in this type of environment that issues such as risk perception and willingness to take risks affect decision making (Wickens et al., 1998). Some examples of the environments in which problem solving and thus naturalistic decision making are necessary could include a country's preparations for an impending natural disaster, a government's response to a bombing within their country, and a pilot's decision to deploy the BRS parachute in a flight emergency. All of these decisions involve multiple issues, and each option has potentially negative consequences. The NDM theory accounts for the cognitive complexity of decision-making, especially decisions made under risk with time constraints, where it is not possible to mentally weigh all of the choices (Wickens et al., 1998). Furthermore, studies have shown performance in aeronautical decision making is not closely related to information processing skills (such as STM) thought to be essential in classic decision strategy (Wickens, Stokes, Barnett, & Davis, 1987). Therefore NDM theory, not traditional or classic theory, will be addressed further in this study, particularly in terms of the most appropriate training methodology.

One example of naturalistic decision-making is Klein's (1989) notion of the Recognition-Primed Decision (RPD). RPD is a model of how experts make quick decisions based on their extensive domain specific knowledge. The idea is that experts recognize a pattern of cues, recall a previous response to the cues, and then implement that course of action. Experts are able to make decisions quickly because of the pattern matching, which is a rapid process by nature. Simon (1987) refers to this process as "intuition". According to Klein (2000), decision makers rely on previous experience to make a decision for the first time. If their choice is successful, experts may reference and use the first decision immediately if confronted with a similar problem. The typical decision is usually made immediately after experienced pilots come up with one option, without waiting to compare it to another alternative. If pilots cannot quickly identify the problem in the situation they may create stories to explain the event and/or mentally predict how the option will be carried out. However, it is possible for the decision maker to follow an analytical process, such as that suggested by classical decision making theory, in certain situations with ample time to compare multiple options.

This theory of RPD is most applicable to situations with time constraints, as with the deployment of the BRS parachute. Although RPD is a decision strategy commonly used by experts, it is not the only theory of NDM, nor is it used by everyone (Klein, 1997). When examining decision making, it is best to understand that many different strategies may be employed by decision makers and the strategy used depends on various factors.

Factors Affecting Decision Making

There are several factors that may affect decision making. The course of action taken, or the decision made, depends on a variety of internal and external factors. External factors might include the amount of time the decision maker has and the number of options that are available, while internal factors involve the individual's personality, motivation, level of experience, and understanding of the consequences of a decision (O'Hare, 1992). Stress and fatigue are also noted to affect decision making (Turner, 1995).

Risk assessment. One particularly important factor is risk awareness. Risk, for the purposes of this study, is defined as the probability that damage will result from a hazard (Wiegmann, & Goh, 2000). Critical decision making tasks, including many in aviation, typically involve risk (Medin & Ross, 1992). Also, decision making and risk management are linked as one behavior in some definitions of CRM. Therefore, pilots who excel at assessing risk and then take the appropriate course of action should fare better than pilots who are less capable of risk assessment. The risk assessment abilities of the pilot population is an issue that requires attention from researchers, as pilots tend to exhibit low levels of risk awareness, and this can lead to greater danger while flying (O'Hare,1990). Pilots are often overconfident in their abilities and many do not fully recognize the magnitude of risks that are present in a situation (Wiegmann & Goh, 2000).

When examining how critical decisions are made, an important characteristic to consider is whether the decision maker tends to be risk averse, risk seeking, or in between the two polar types. Both younger pilots and more experienced pilots seem *less* conservative, or more willing to take risks, when making decisions (Driskill, Weissmuller, Quebe, & Hand, 1998). Similar findings are reported in a study assessing pilots' subjective "level of comfort" when flying. That is, those pilots reporting a high level of comfort are the least risk averse (Driskill, Weismuller, Quebe, Hand, Dittmar, & Hunter, 1997). More experienced pilots have demonstrated better decision making skills, most likely because they are able to combine pieces of the situation together as a whole, unlike novices who lack relevant knowledge (Klein, 1998; Chase & Simon, 1973). Pilots with more experience have also been found to make better decisions under stress, when compared to less experienced pilots (Stokes, Kemper, & Marsh, 1992). Domain specific knowledge is believed to be essential in the diagnostic stage of decision making, because it reduces the mental workload by requiring less information to be kept in working memory (Goh & Weigman, 2002).

Factors affecting the ejection decision. Goodman (1998) cites similar factors which affect the ejection decision, based on his own experience with military pilots and a review of relevant literature. Some of the factors were response time, pressing (where a desire to be successful supersedes rational decision making), situational awareness or lack thereof, over concentration, the stigma associated with losing an aircraft, attempting to overcome the problem for an excessive amount of time, complacency, behavioral inaction or "freezing up", and temporal distortion.

Sturgeon (1988) also cites reasons why pilots may delay the ejection decision. The factors include: reaction times; personality; temporal distortion; complacency; lack of training; desire to move plane away from populated areas; problems with technique; fatigue; and stress. Turner (1995) offers similar reasons for ejection delays.

Sturgeon (1988) notes the influence of the cockpit environment and physiological state of the pilot on the processing and judgment capabilities of the pilot. Too much or not enough sensory input from the environment can have a negative impact on decision formulation, or *processing*. The quality of the input must also be considered, as pilots tend to ignore some sensory input because it is often unreliable in flight, (e.g. the Coriolis illusion is caused by vestibular disturbance, where the pilot falsely perceives a change in direction although the plane maintains the same course). The mode the information is received in may have an impact on sensory awareness. For example, if one mode (e.g. auditory channel) is overloaded, pilots may have better perception if the salient information is presented through another channel (e.g. visually). Processing capability also varies among individuals and across situations. The arousal level, or the physiological manifestations of stress, is affected by situational factors, e.g. is the matter life or death, how much time the pilot has to make a decision, how familiar is the situation, etc. Individual factors could also play a role in determining pilot arousal – some pilots may be affected differently by the same amount of stress. For example, according to the Yerkes-Dodson Law, performance decrements caused by stress/increased arousal appear more rapidly with complex tasks rather than simple ones. An inverted U exists between performance and arousal in general, where too little or too much arousal can have a negative impact on performance. The extent of impact depends on the situation and the amount of stress. In certain situations (e.g. emergencies) where the consequences of decisions are critical and events often unfold rapidly, stress can overwhelm the decision maker and have a detrimental effect on the quality of the decision. An experienced or well-trained pilot may be able to react automatically to the situation, thereby avoiding the negative impacts of stress on decision making.

A false hypothesis, or a decision error made because sensory input was incorrectly interpreted, is more likely to occur under one or more of the following conditions: the pilot has certain expectancies regarding the incidence of events; arousal levels are altered or were recently very high; attention is distracted; or false assumptions have been maintained by the pilot for an extended amount of time (Sturgeon, 1988).

Ejection Seats and Decision Making

Many researchers have examined the human factors issues associated with ejection seats. This includes training pilot decision making regarding the use of ejection the devices. For example, Callaghan & Irwin (2001) studied factors influencing a pilot's decision to eject from an aircraft. These authors postulated a lowered ejection decision height would result in pilots following the "prescribed" ejection height; thereby allowing further aircraft maneuvering to result in a safer ejection position. The ejection decision height did not appear to influence the accuracy of the decision to eject, but did affect the bias for ejecting. A lower height resulted in a greater bias for ejecting. Experienced pilots were also noted to be more decisive in making their judgments than less experienced pilots. Another factor noted to influence decision-making is the nature of the emergency -- unfamiliar situations do not allow for schema-based decision making.

Now that the relevant KSAOs have been discussed, the training method selected for this study will be identified next.

Training Method

As noted previously, the main skills considered to be important in deciding whether or not to deploy the BRS parachute are situational awareness, decision-making and risk assessment, all of which are considered part of SRM. Knowledge of which situations require the parachute is also essential in using the parachute successfully.

A training method must be selected before the training intervention is planned in full detail. There are many basic training methods to choose from, including but not limited to – lecture, audiovisual techniques, self-directed learning, simulation, and many others (Noe, 1999).

Aviation training involves complex interactions between humans and technology. The aviator must possess all of the necessary aviation knowledge and skills before flying in such a complex environment alone (Oser, 1999). Indeed, effective human performance can mean the difference between life and death. The operator must constantly assess the environment, observe any salient signals or patterns, and then make immediate decisions to react appropriately to the environment (Oser, 1999). Competencies may specifically include aeronautical decision-making, situational awareness, critical thinking, and a number of other high order skills. As noted by Oser Cannon-Bowers, Salas, & Dwyer (1999), “an effective training method will facilitate the ability of participants to develop these necessary competencies.”

Klein (2000) suggests due to the vagueness of "training decision making skills", training programs designed to help pilots make better decisions should focus on specific reactions for specific situations. Scenario-based training (SBT), one method recently used in aviation and other fields, utilizes practice and feedback with specific scenarios to improve performance and may lead to improved decision making as a result of the training (Oser, et al. 1999). SBT also utilizes extensive practice to automate complex tasks (Oser, 1999). An aviation example is to train pilots to react appropriately in an emergency situation based on information from their surroundings. The scenario events, e.g. icing on aircraft wings, for training such reactions would allow pilots to practice making decisions while under extensive stress, thus automating the process over time. This study seeks to examine the use of a specific training technique, scenario-based training, to accomplish this task.

Scenario-Based Training Approach

Traditional Training Methods

In traditional aviation training methods, trainees are lectured, learning facts and procedures first through memorization. Trainees are then evaluated with written tests and may later practice in a simulator, typically performing tasks one at a time (Karp, 2001). According to interviews with subject matter experts, aviation training traditionally involves practicing training tasks one at a time, such as certain maneuvers, without integrating the individual tasks into a realistic flight. The training process has changed minimally over time, despite the increased complexity of aircraft and technology of simulators (Karp, 2001). In some cases, simulation may be used very little, if at all, before the students plunge into flight time. However, simulation devices are recommended for use in flight instruction, as improved technology allows important features, such as weather, to be recreated realistically (Fowlkes, Dwyer, Oser & Salas, 1998). Additionally, simulation devices provide a more comprehensive training experience, as they allow pilots to fly in scenarios that might be too unusual or dangerous for novice pilots to practice in real flight. Certain tasks are more effectively trained in a simulation device than others, such as takeoff and landing (Hays, Jacobs, Prince, & Salas, 1992). Therefore, it seems aviation training, including general aviation, should incorporate simulation more frequently and/or earlier in the curriculum. A training method specifically designed for use with simulation should be examined to address these concerns.

Line-Oriented Flight Training (LOFT) Methods

According to Lauber and Foushee (1981), LOFT refers to aircrew training which involves the simulation of realistic full mission situations. These situations, or scenarios, consist of typical daily operations for the airline and are often developed from accident reports to enhance realism. Realistic problems and emergencies are introduced during the scenarios to train correct flight deck management techniques. After a session is completed, a comprehensive debriefing begins with a self analysis by the crew. This is followed by the LOFT coordinator's debriefing. Voice and video recorders, along with written notes are recommended for the debriefing session.

LOFT is believed to have the potential to significantly impact aviation safety through improved training and validation of operational procedures. Indeed, it is approved for use instead of the usual semi-annual proficiency checks in aviation, although certain conditions must be met. Although LOFT is used to enhance decision making in aviation, emphasis is placed on situations which involve communications, management and leadership. These goals differ somewhat from those of a single pilot in general aviation, so a slightly different method was used in this study and is discussed next.

Scenario-Based Training Methods

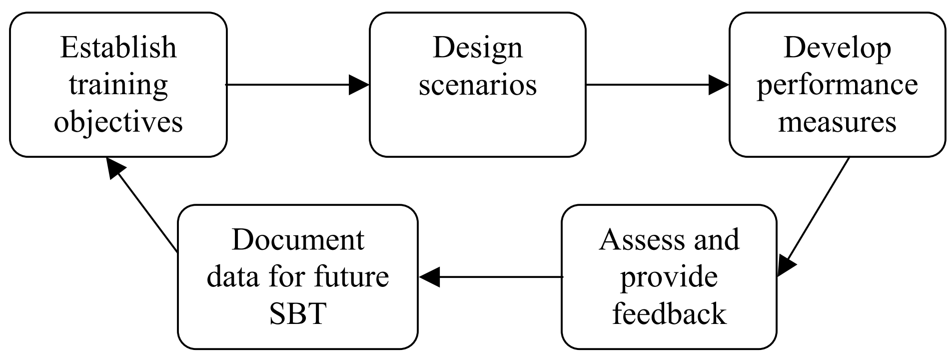
Scenario-based training is a similar and closely related form of training which uses simulated scenarios to train tasks. Cannon-Bowers, Burns, Salas, & Pruitt, (1998) describe SBT as training that focuses on well-planned exercises, with feedback provided to trainees based on their responses to simulated cues comparable to those in the actual work environment. SBT differs significantly from traditional methods, as the scenarios comprise the entire curriculum from the beginning (Cannon-Bowers, et al.1998). Pilots learn by performing tasks and making improvements based on recommendations from the instructor. A variety of tasks are presented to the trainee to quickly provide an inexperienced operator with practice in decision making and other related skills. Practice is essential in decision making, as pilots often have a short amount of time to make critical decisions, so it is best if the process is automated as much as possible (Turner, 1995). A wide range of operators may benefit from SBT, from novices without any specialized knowledge to expert users who wish to familiarize themselves with a new product (Loftin, Wang & Baffes, 1989).

SBT further deviates from traditional methods. Traditionally, pilots practice a single flight task until that task is mastered, generally with the instructor guiding them through the training. The individual tasks are not immediately integrated and practiced as one complete process. This method may be appropriate for the mastery of some simple tasks, but the complex interactions between pilots, instrumentation and ATC that arise when flying a TAA may require an alternative training method (Oser, 1999). In contrast, SBT allows pilots to practice many tasks simultaneously within a scenario of simulated flight and pilots do not receive instructions during flight. Thus, pilots are quickly taught to make autonomous decisions, as they go through the entire scenario on their own and receive feedback about their performance after the scenario is completed. With this method pilots are trained to understand the interactions from the beginning, and thus have more time to develop a variety of skills.

The SBT Process

A crucial aspect of SBT is the *structure* or the linking together of a number of important aspects: learning objectives, scenario events, performance measures, and feedback (Oser et al., 1999). An integrative approach is used to ensure that "all aspects of scenario design, development, implementation, and analysis" are linked (Oser, 1999). This systematic approach ensures KSAOs are addressed. As described in Oser et al. (1999), and as shown in Figure 2, the first step is to identify learning or training objectives (i.e., the knowledge and skills necessary to perform a specific task). Training objectives are usually based on a noted need for improvement, a requirement for new skills, and/or the introduction of unfamiliar equipment. Next, scenarios are designed based on the training objectives. The scenarios consist of a series of events that give learners the opportunity to practice using the essential knowledge and skills and are carefully planned to ensure all training objectives are met (Cannon-Bowers et al., 1998). For example, one event might be an aircraft losing all engines during flight. Next, performance measures are developed which allow an instructor to assess the degree to which a trainee is demonstrating the correct knowledge and skills (Oser et al., 1999). Finally, after a trainee experiences/performs the scenario and the instructor assesses performance, the instructor gives the trainee feedback on his/her performance (Oser et al., 1999). The feedback is specific and targeted, and enables the trainee to know exactly what was right and what needs to be corrected, so performance of future scenarios is improved. In sum, the SBT approach offers a structured approach while providing pilots the opportunity to "train the way they fly."

Figure 2. Scenario Based Training Model, adapted from Oser et al., 1999.



SBT Research

Although research on SBT is limited, studies have found positive results using SBT. Results have indicated learners typically believe SBT is an effective training method. For example, stressful and realistic scenarios were presented in a training program for firearm safety with the goal of increasing mental readiness of probation officers (Scharr, 2002). In a post-training evaluation questionnaire, 97% of the respondents felt the training was effective "to a great extent." Additionally, Lowry (2000) found that probation officers trained with SBT were much more likely to rate the training as excellent compared to officers trained with other methods.

Training cost savings have also been reported. Using a method similar to SBT, Stewart, Dohme & Nullmeyer (2002) examined the effectiveness of pilot training for rotary wing aircraft. Training times and costs were reduced as a result, and the transfer of training ratio was found acceptable. Cannon-Bowers et al. (1998) also claims training time and costs are reduced as a result of including only the events which exercise the targeted skills.

Scenario based training has been used in the medical field to allow doctors to practice techniques such as sigmoidoscopy (Kneebone et al., 2003). Using a simulated patient, participants were able to practice their procedural skills without risking harm to a live patient. Scenario-based assessments showed an improvement in performance due to simulation-based practice.

SBT is believed to be a very effective method for training with simulation devices. Critical aspects of the environment are accurately reproduced in a high fidelity simulation device, so high order skills and procedural knowledge may be developed with practice and corrective feedback (Oser et al., 1999). In a related study, a network of simulation devices was used to train military service personnel (e.g., U.S. National Guard, U.S. Marine Corps, U.S. Army and U.S. Air Force) around the world in a joint military exercise. In this exercise, Dwyer, Oser, Salas, & Fowlkes (1999) tested the use of performance measurement tools designed to specifically link to events and, thus, learning objectives. Although only case studies were used, the results indicated the SBT approach was a success.

With its emphasis on reproducing important characteristics of the operational environment and its use of simulated environments, SBT is thought to be particularly useful, if not essential, for training tasks within complex environments--such as tasks that occur within a cockpit (Oser et al., 1999). Indeed, initial work indicates SBT is an effective training method for military aircrew (Oser et al., 1999).

Researchers have also begun to assess whether a variation of SBT will be a successful training method specifically for CRM. Hedge, Borman, & Hanson (1996) utilized video based training in place of more expensive simulation training or traditional lectures. Video-taped scenarios were presented to participants, who then discussed the actions taken within the scenarios. This form of training has been well received, although it does not appear to have been validated.

Based on Oser et al. (1999) and other related work, the scenario-based training method seems an appropriate method for training pilots to effectively use the BRS parachute system. SBT has been found to be an effective method, reducing response time, enhancing high order and procedural skills, reducing training costs and time, and has been rated as highly effective by former SBT trainees (Bowers & Morgan, 1991; Scharr, 2002; Lowry, 2000; Stewart et al., 2002). SBT has also been used in a variety of domains with apparently successful results. Further research on SBT effectiveness, however, would enable training designers to take maximal advantage of the method while avoiding pitfalls. Additionally, the BRS parachute should be effectively trained using simulation because of the complexity involved in the decision to release and because it is an alternative means to land an aircraft. Also, it is not practical to use the BRS parachute when training in a real aircraft. Therefore, a training tool such as a flight simulation device is necessary to precisely recreate the important features of the environment and for deploying the BRS parachute.

Purpose

At the same time, more research is needed regarding the efficacy of the SBT approach. The results of this study will:

1. Design and validate a training intervention that may be used to ensure pilots will properly use the BRS parachute.
2. Compare the accuracy of the decisions made by pilots, prior to and after training, regarding the deployment of the BRS to the decisions pilots should make according to subject matter experts.
3. Provide additional data on the effectiveness of SBT as a training methodology.
4. Provide insight regarding future training needs and/or equipment design to facilitate effective parachute use.

Hypothesis

Based on relevant research, scenario-based training will prove to be an effective method for training the BRS parachute and will effectively train the necessary KSAs, such as decision-making and risk awareness. Participants trained via SBT on use of the BRS parachute will perform significantly better on a variety of measures than the participants in the control condition.

Hypothesis 1: Participants in the SBT condition will perform more effective parachute use behaviors than will the participants in the control condition when tested with emergencies within scenarios in a simulation device. Participants in the SBT condition will also perform more effectively than participants in the control condition for the following behaviors:

Subhypothesis 1a: “Controlled landing/BRS decision”

Subhypothesis 1b: “Looked for place to land”

Subhypothesis 1c: “BRS when necessary”

Subhypothesis 1d: “Correct BRS use”

Subhypothesis 1e: “SRM”

Subhypothesis 1f: “Overall performance”

Subhypothesis 1g: “Frequency did not crash”

Subhypothesis 1h: “BRS timing”

Subhypothesis 1i: “BRS altitude”

Subhypothesis 1j: “BRS knots”

Subhypothesis 1k: “Maintains control of aircraft”

Subhypothesis 1l: “Refers to checklist”

Subhypothesis 1m: “Follows checklist”

Subhypothesis 1n: “Contacts ATC”

Subhypothesis 1o: “Declares emergency”

Subhypothesis 1p: “Diverts”

Hypothesis 2: Participants in the SBT condition will achieve significantly higher scores on the knowledge test than the participants in the control condition.

Subhypothesis 2a: Participants in the SBT condition will achieve significantly higher scores on the BRS portion of the knowledge test than will participants in the control condition.

Subhypothesis 2b: Participants in the SBT condition will achieve significantly higher scores on the SRM portion of the knowledge test than will participants in the control condition.

Hypothesis 3: Participants in the SBT condition will have significantly higher self –efficacy, as measured by the Self-Efficacy Questionnaire (SEQ), regarding the use of the parachute than will the participants in the control condition.

Subhypothesis 3a: Participants in the SBT condition will have significantly higher ratings of self-efficacy than the control condition on the BRS portion of the Self-Efficacy Questionnaire.

Subhypothesis 3b: Participants in the SBT condition will have significantly higher ratings of self-efficacy than the control condition on the SRM portion of the Self-Efficacy Questionnaire.

Hypothesis 4: Participants in the SBT condition will exhibit significantly lower levels of perceived workload than participants in the control condition.

METHOD

Independent Variable

A scenario-based training intervention was developed to instruct pilots in the knowledge and skills necessary for effective use of the parachute. The intervention described the "what, how, and when" of parachute use and gave trainees the opportunity to experience many realistic scenarios pertaining to parachute use and make decisions regarding using or not using the parachute. Trainees received feedback on their decisions regarding the use of the parachute. A traditional training group acted as the control condition, and both conditions are described in further detail next.

Scenario-based training intervention

Participants in the scenario-based training condition discussed emergency procedures in detail with the experimenter, including when it is safe to use the BRS and the severity of an emergency which warrants using the BRS. Participants then flew a simulation device and faced extreme and unusual events within scenarios during the cruise phase of flight, and had to make decisions regarding the appropriate course of action to take when a critical problem arose. Four different scenarios were presented to trainees, with a total of ten different emergency events occurring in all scenarios. The events in the scenarios consisted of a variety of emergency situations that could arise when flying an airplane, such as engine failure and icing. The emergency situations varied in terms of consequence and severity and arose at different times within the scenarios -- ranging from 5 to 30 minutes from the beginning of flight. Subject matter experts determined in advance whether each situation would merit the deployment of the BRS parachute. Three of the ten events most likely required the use of the parachute. Participants received feedback from a flight instructor at the end of each scenario. The feedback was based on the SMEs assessment of the appropriate action the pilot should take within the scenario. Training scenarios used in this study may be found in Appendix A. The total scenario-based training session time for the parachute was about three hours.

Traditional training intervention

In the aviation industry, training programs rarely cover the BRS parachute, as only two makes of planes currently come standard with this equipment. At the time of this study, the programs that do include the BRS appear to offer a minimal amount of training, typically just describing the sequence of steps to use the BRS and providing little instruction on *when* to use it. Pilots might be told there are several situations that might warrant using the BRS parachute, such as engine failure at night or over hostile terrain, but more specific information is not given. The traditional training condition was similar to the current methods and therefore represented traditional aviation training for the BRS parachute. Participants in this group received information from the instructor regarding the procedure for using the BRS parachute. This information was based on information currently provided by an industry training program. Emergency procedures were discussed as with the SBT group, including some scenarios in which BRS use is appropriate and to refer to checklists in an emergency. Participants were given a computer-based demonstration of the procedure for deploying the parachute and had the opportunity to practice the procedure both via computer and within the simulation device. Pilots then read industry generated packets of information about the BRS parachute and SRM. Next, participants in this condition flew the same training scenarios as the SBT group in the simulation device, but without emergency events and feedback from an instructor. However, training time was about equivalent for both groups. Participants in the traditional condition spent about as much time reading and using Computer-Based Training (CBT) as the SBT condition spent receiving practice and feedback with emergencies. The total time for the traditional training condition was about three hours.

Dependent Variables

The effect of the training intervention on the pilot performance was assessed with a number of criteria. These correspond to Kraiger, Ford, & Salas (1993) multiple measures of learning. These include pilot performance, knowledge acquired, self-efficacy, and perceptions of stress.

Pilot Performance

Two SMEs assessed pilot performance using rating scales in both the pre and posttest scenarios. The raters were blind to condition. Both raters watched the pre and posttest (either in person or by video), then rated each participant using the performance measures and rating scales found in Appendix B. The rating scales included specific behaviors which are believed to demonstrate desired KSA's for parachute use, such as pilot situational awareness, decision making skills and BRS knowledge. These behaviors were rated using either a five point Likert scale or a yes or no categorization. A five on the five point scale represents a very appropriate behavior or response was made by the participant, a three represents a moderately inappropriate behavior, and a one represents a very inappropriate behavior. For example, if the appropriate behavior was pulling the parachute, a participant received a five if he or she uses the correct procedure to release the parachute. Each performance measure was used at least twice in the posttest to enhance reliability of the assessment. Equivalent performance measures were averaged for each participant and these total averages were used in data analysis. After each scenario, raters met to discuss their assigned ratings. The two raters then developed one consensus rating for each performance measure at this meeting.

Knowledge Test

Knowledge of the parachute was assessed with an averaged score on a knowledge test (see Appendix C). The parachute test was 10 questions long. Seven questions ask pilots how, when and where it is appropriate to use the BRS parachute. Three questions assess their knowledge of SRM. A reliability analysis was performed for the knowledge test to determine if the questions should be averaged together. Using Cronbach's Alpha to compute the reliability, BRS related questions on the SEQ had sufficient internal consistency to be analyzed together (r = .70). However, using Cronbach's alpha on SRM related questions resulted in a negative number, which violates reliability model assumptions. This was probably due to the different variances for each of the SRM questions. Therefore, Guttman's split-half reliability coefficient was used for the reliability analysis of SRM questions. Internal consistency was acceptable with this analysis (r = .66). The knowledge test was thus divided into two sections for data analysis. An averaged score for BRS related questions was used in analysis along with an averaged score for SRM related self-efficacy questions.

Self-efficacy Questionnaire

Self-efficacy is an individual's belief in their ability to succeed at a specific task (Bandura, 2000) and is an important consideration in training evaluations (Gist, 1989). A high level of self-efficacy is linked to superior performance (Bandura, 1997; Bandura, 2000). According to Locke & Latham (1990), the mean correlation for self-efficacy and performance goal setting is also estimated to be fairly high across studies (r= .39). Participant's self-efficacy for using the BRS parachute and SRM was assessed with a ten question survey, using a Likert scale (see Appendix D). The scale for this study was adapted from a validated scale for self-efficacy (Riggs, 1989). Seven of the items on the questionnaire are related to self-efficacy for the BRS parachute in emergency situations. The three remaining items assess self-efficacy for SRM knowledge and abilities. A reliability analysis was run on the Self-efficacy Questionnaire (SEQ) to determine if the questions should be averaged together. Using Cronbach's Alpha to compute the reliability, BRS related questions on the SEQ had sufficient internal consistency to be analyzed together (r = .70). SRM related questions also had high internal consistency (r = .78). Therefore, the SEQ was divided into two sections for analysis. An averaged score for BRS related self-efficacy measures was used along with an averaged score for SRM related self-efficacy measures.

Workload

High-risk events are often accompanied by stress. Since workload is one stressor (Wickens, Gordon, & Liu, 1998), subjective measures of workload will help quantify whether participants perceive they have a lower workload in the performance tests after having had the parachute training. The idea is if they perceive less workload, they will have more mental resources available to make effective decisions regarding the parachute. Perceptions of workload were assessed using an abbreviated version of the NASA TLX after the pre and posttest in this experiment (see Appendix E). The NASA TLX is a validated multidimensional workload measure given to the participant after the pre and posttests (Hart & Staveland, 1988). It measures six dimensions of workload: mental demand; physical demand; temporal demand; performance; effort; and frustration. A reliability analysis was performed on all questions of the NASA TLX to determine if the questions should be averaged together. Using Cronbach's Alpha to compute the reliability, questions on the NASA TLX (r = .80) had sufficient internal consistency to be analyzed together. Therefore, an averaged score for the participant's pre and post TLX ratings was used for data analysis.

In summary, the multiple measures of behavior and knowledge as well as the subjective measures of pilot self-efficacy and workload were considered together to determine the effectiveness of the parachute training intervention.

Participants

Thirty-six participants were recruited from the ERAU student pilot population. All pilots had at least a private license and instrument rating with 100 - 300 hours of total flight time. Thirty-five of the participants were male; only one participant was female. The mean age for all participants was 20.8. Participants had an average of 3.2 years of piloting experience and 199 flight hours. No significant differences were found between groups for these demographic variables.

Apparatus

An Elite™ flight simulation device was used in this experiment. This represented a traditional general aviation cockpit. Microsoft Flight Sim and Microsoft Flight Sim for Instructors software were used in conjunction with the Elite™ flight simulation device for this experiment.

Materials

A demographics questionnaire will request background information from participants such as how many flight hours they have, whether they have previously flown a plane or a simulation device equipped with a parachute, etc. (see Appendix F).

Procedure

Participants were randomly divided into two groups. Random assignment was used to control for any differences in age, experience, training and other characteristics. Eighteen pilots were assigned to the experimental/SBT group. These pilots received the in-depth parachute scenario-based training intervention. Eighteen other pilots received parachute training via traditional BRS training methods (i.e. computer-based training).

All participants signed a consent form when they first arrived for the experiment (see Appendix G). Next, participants planned for the training pretest flight using sectionals and other information provided to them. Participants were then shown the simulation device and where the handle for the BRS parachute was located in the cockpit. They received a brief description of the parachute and were instructed to treat the simulation device like a real plane. Participants ran through the before take-off checklist and then began their first flight in the Elite™ flight simulation device. At 600 ft AGL after takeoff, engine failure occurred in the scenario and participants had to respond quickly to this emergency. The best response to this particular emergency was to use the parachute immediately. Performance assessments (provided by SMEs) were used to rate the participants performance/response during the pre-test scenario. Two raters, also SMEs, assessed pilot performance in person or via videotape. Altogether the pretest lasted about 20 minutes, which included 15 minutes for pre-flight planning and 5 minutes for the scenario.

Following the pretest, participants completed the NASA-TLX and demographics questionnaire. Participants were not given feedback about their performance until after the forms were completed. Next, participants received either SBT or traditional training for the parachute described previously. Upon completion of the parachute training, participants were dismissed for the day. The following day all participants received a post training performance test. This posttest involved flying the simulation device for three scenarios with seven emergency events. The scenarios and emergencies were similar to the scenarios in the training condition with some variations. These scenarios were designed to demonstrate their parachute use skills and knowledge. During the performance test, the participants were not given feedback and did not receive any instruction. Two of the events required the use of the BRS parachute to avoid terrain collision. The same trained raters used for the pretest assessed performance during the posttest scenarios. Both raters assessed the pilot's performance separately for each scenario, then held a brief meeting to establish a consensus rating for each measure.

All pilots completed the NASA TLX for a second time after the performance posttest. The self-efficacy questionnaire was then administered to pilots and was followed by the BRS and SRM knowledge test.

RESULTS

The purpose of this study was to examine the effects of two different training methods on BRS use and knowledge.

Bivariate correlations were performed on all dependent variables and a correlation matrix is presented in Table 1. Many of the performance measures were significantly correlated at the .05 level. Several performance measures were also correlated with one of the demographics questions, number of flight hours. A few performance measures are also correlated with the SEQ and knowledge test. Therefore, some of the variables appeared to be measuring similar items. This finding is appropriate as many involve the BRS parachute or SRM, yet are not assessing the exact same behaviors or attitudes.

Table 1. **Correlation Matrix**

		Control Landing	BRS use	BRS timing	BRS altitude	BRS knots	SRM	Overall	Freq not crash	Control aircraft	Refer to checklist	Follow checklist	Contact ATC	Declare emer	Divert	TLX Pretest	TLX Posttest	SEQ BRS	SEQ SRM	Know test BRS
BRS timing	r	0.502**	0.547**																	
	N	35	35																	
BRS altitude	r	0.238	0.579**	0.500**																
	N	34	34	34																
BRS knots	r	-0.023	0.465**	-0.022	-0.024															
	N	33	33	33	33															
SRM	r	0.474**	0.027	0.504**	0.143	-0.189														
	N	35	35	34	33	32														
Overall	r	0.591**	0.352*	0.670**	0.489**	-0.230	0.656**													
	N	36	36	35	34	33	35													
Freq not crash	r	0.470**	0.626**	0.583**	0.415*	0.085	0.287	0.483**												
	N	36	36	35	34	33	35	36												
Control aircraft	r	0.066	0.261	0.191	0.294	0.216	0.361*	0.370*	0.390*											
	N	36	36	35	34	33	35	36	36											
Refer to checklist	r	0.001	-0.337*	0.025	-0.004	-0.220	0.153	0.334*	-0.130	0.176										
	N	36	36	35	34	33	35	36	36	36										
Follow checklist	r	0.018	-0.219	0.064	-0.010	-0.169	0.201	0.331*	-0.122	0.200	0.851**									
	N	36	36	35	34	33	35	36	36	36	36									
Contact ATC	r	0.356*	-0.036	0.237	0.078	-0.098	0.513**	0.534**	0.025	0.315	0.349*	0.299								
	N	36	36	35	34	33	35	36	36	36	36	36								
Declare emer	r	0.356*	0.021	0.280	0.177	0.020	0.508**	0.595**	0.069	0.293	0.372*	0.364*	0.758**							
	N	36	36	35	34	33	35	36	36	36	36	36	36							
Divert	r	0.191	-0.284	0.095	-0.028	-0.209	0.354*	0.388*	-0.072	-0.037	0.398*	0.430**	0.310	0.377*						
	N	36	36	35	34	33	35	36	36	36	36	36	36	36						
TLX Pretest	r	0.110	0.095	0.020	0.158	-0.222	-0.084	0.002	0.122	0.180	-0.129	-0.140	-0.012	-0.092	0.060					
	N	36	36	35	34	33	35	36	36	36	36	36	36	36	36					
TLX Posttest	r	-0.056	-0.129	-0.084	-0.182	-0.020	-0.186	-0.160	-0.074	-0.089	-0.183	-0.236	0.127	0.200	0.179	0.508**				
	N	36	36	35	34	33	35	36	36	36	36	36	36	36	36	36				
SEQ BRS	r	0.044	-0.114	0.051	0.089	-0.177	0.247	0.200	-0.039	-0.014	0.188	0.093	0.318	0.376	0.126	-0.081	0.115			
	N	36	36	35	34	33	35	36	36	36	36	36	36	36	36	36	36			
SEQ SRM	r	0.264	-0.115	0.012	0.210	-0.182	0.270	0.153	-0.079	0.021	0.068	0.069	0.373*	0.437**	0.199	-0.112	-0.008	0.405*		
	N	36	36	35	34	33	35	36	36	36	36	36	36	36	36	36	36	36		
Know test BRS	r	0.166	-0.166	-0.049	-0.170	-0.295	0.161	-0.049	-0.093	-0.097	-0.006	-0.088	0.062	-0.154	0.006	0.000	-0.163	-0.312	0.033	
	N	36	36	35	34	33	35	36	36	36	36	36	36	36	36	36	36	36	36	
Know test SRM	r	0.098	0.178	0.322	0.205	-0.176	0.423*	0.399*	0.040	0.237	-0.051	0.129	0.465**	0.368*	0.110	0.166	0.114	0.095	0.131	0.015
	N	36	36	35	34	33	35	36	36	36	36	36	36	36	36	36	36	36	36	36

*. Correlation is significant at the 0.05 level (2-tailed). **. Correlation is significant at the 0.01 level (2-tailed).

The multiple hypotheses for this study were analyzed using one of two methods, either a two-way ANOVA or a one-tailed independent samples t-test. Due to the scale used to assess performance and the small number of participants who selected to use the parachute in the pretest, there was no pretest performance data available for some of the measures. Condition was a between subjects independent variable for all measures in this study. For the measures with pre and posttest data, session was included as a within subjects independent variable. One-tailed tests were used because of the directional hypotheses. The method of analysis was selected mainly based on whether or not pretest data was available and is described below.

Pilot Performance

Hypothesis 1: Participants in the SBT condition will perform more effective parachute use behaviors than will the participants in the control condition when tested with emergencies within scenarios in a simulation device. This hypothesis was decomposed further in order to specifically examine the individual BRS related behaviors. All performance measures may be found in Appendix B.

It was hypothesized that participants in the SBT condition would perform significantly better than participants in the CBT condition on the performance measure “Pilot makes appropriate controlled landing/BRS decision” (subhypothesis 1a). The mean and standard deviation for the pre and posttest performance measures for both conditions is shown in Table 2. Figure 3 depicts the mean and standard error of the mean for session and condition. Figures 4 and 5 contain box plots for the pre and posttest. The box plots illustrate the median, interquartile range (the difference between the 75th percentile and the 25th percentile), and minimum and maximum rating. This subhypothesis was analyzed with a two-way ANOVA. As shown in Table 3, a main effect was found for session, $f(1, 32) = 6.15$, $p = .02$. Thus, session was significantly related to ratings of "controlled landing/BRS decision". An eta squared of .14 indicates that 14% of the variability in controlled landing/BRS decision is related to differences in session. The means of the posttest were higher than the means of the pretest across condition. A significant main effect was also found for condition, $f(1, 32) = 4.29$, $p = .05$. Thus, condition was significantly related to ratings of "controlled landing/BRS decision". An eta squared of .07 indicates that 7% of the variability in "controlled landing/BRS decision" is related to differences in condition. The SBT group had an overall higher mean across time (pre-post). However, the interaction for this performance measure was not significant, $f(1, 32) = 3.94$, $p = .06$. Therefore, this subhypothesis was not supported.

***Table 2.* Descriptive Statistics**

Table 2

Means and Standard Deviations of Performance Data for Pretest and Posttest Ratings on BRS and SRM Measures

Performance Measure	Test	*N*		*M*		*SD*	
		SBT	CBT	SBT	CBT	SBT	CBT
Controlled Landing	Pre	17	17	2.88	2.94	.86	.74
	Post	17	17	3.94	3.06	.68	1.22
Correct BRS use	Pre	14	15	1.14	1.33	.53	.90
	Post	14	15	3.77	3.81	1.23	1.29
SRM	Pre	16	17	1.00	1.00	.00	.00
	Post	16	17	3.19	1.11	1.18	.24
Overall Performance	Pre	17	17	2.59	2.71	.94	.99
	Post	17	17	3.86	3.00	.45	.58
Frequency did not crash	Pre	18	18	64.71	52.94	49.26	51.45
	Post	18	18	90.52	78.94	15.13	22.57

***Table 3.* ANOVA**

Table 3

Two way Analysis of Variance for "Pilot Makes Appropriate Controlled Landing/BRS Decision" Performance Measure

Source	*Df*	F	η^2	Power	*p*
Session (S)	1	6.15	.14	.67	.02
Group (G)	1	4.29	.07	.52	.05
S × G	1	3.94	.09	.49	.06
S within-group error	32	(.96)			

Note. Values enclosed in parentheses represent mean square errors.

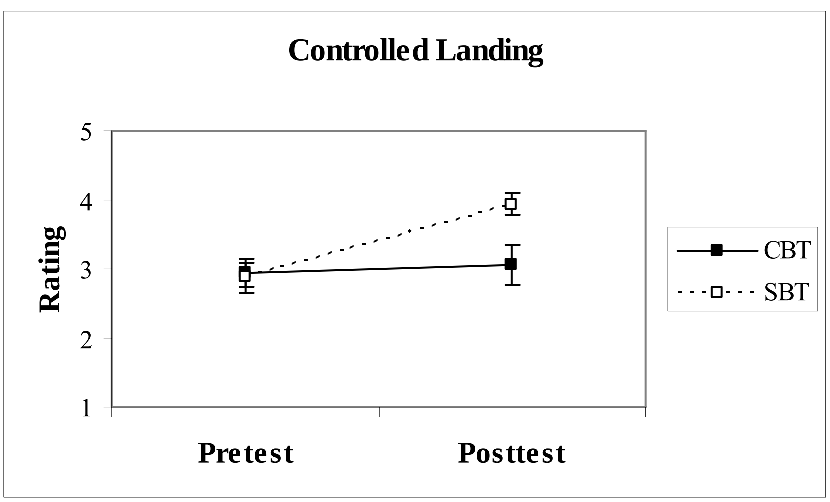


Figure 3. Mean and standard error of the mean for the performance measure "controlled landing/BRS decision", with rating as a function of session and condition.

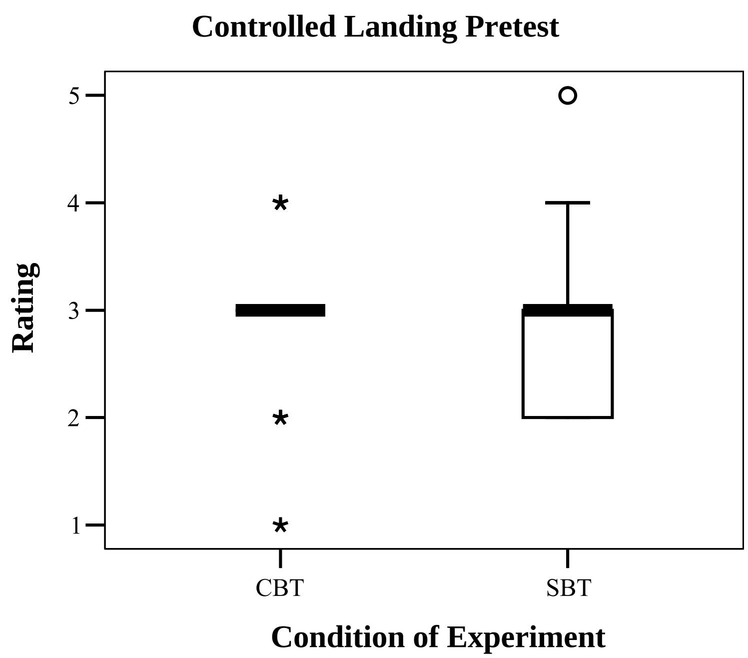


Figure 4. Box plot for the performance measure “controlled landing/BRS decision”, with pretest rating as a function of condition.

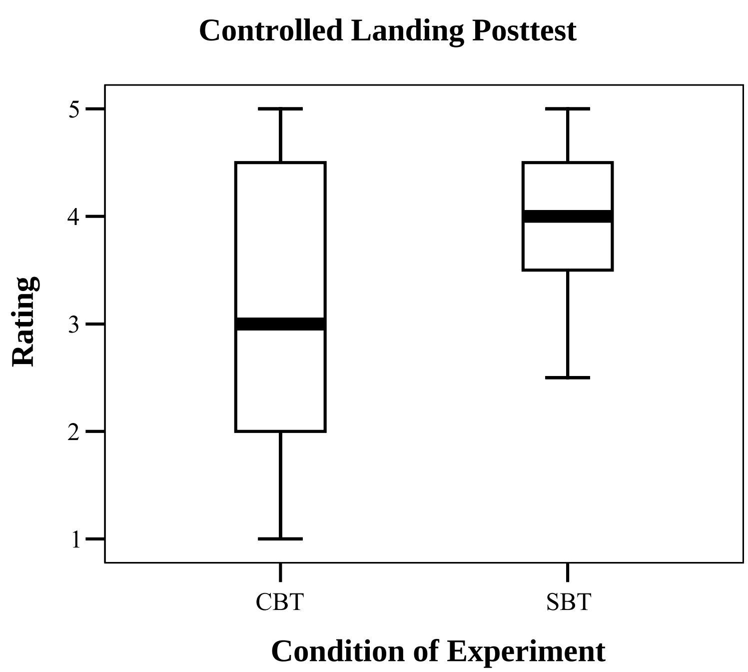


Figure 5. Box plot for the performance measure “controlled landing/BRS decision”, with pretest rating as a function of condition.

The performance measure “Pilot makes appropriate controlled landing/BRS decision” was further decomposed to provide a more detailed level of analysis and address the two major concerns with the BRS parachute: failing to look for a place to land before deploying the parachute and neglecting to use the parachute when necessary. The “controlled landing” performance measure was therefore subdivided into “Percent looked for a place to land before using the BRS parachute” (subhypothesis 1b) and “Percent used parachute when necessary” (subhypothesis 1c). Again, the SBT condition was expected to perform significantly better than the CBT condition on these performance measures. Pretest data was not available for these measures. The means and standard deviations for the posttest data are shown for both measures in Table 4.

Figure 6 depicts the posttest mean and standard error of the mean for the “Percent looked for a place to land before using the BRS parachute” measure (subhypothesis 1b). Figure 7 contains a box plot for both conditions. This subhypothesis was analyzed with an independent samples t-test. As shown in Table 5, a significant difference was found between groups, $t(31) = 2.26$, $p = .02$. Thus, condition was significantly related to ratings of “looked for a place to land”. An eta squared of .15 indicates that 15% of the variability in “looked for a place to land” is related to differences in condition. The mean of the SBT condition was higher than the mean of the CBT condition. Therefore, the SBT group performed significantly better and this subhypothesis was supported.

***Table 4.* Means and Standard Deviations**

Table 4

Means and Standard Deviations for Measures Related to BRS Decision

	n		*M*		*SD*		*St Error Mean*	
	SBT	CBT	SBT	CBT	SBT	CBT	SBT	CBT
Looked for place to land	15	16	90.00	68.75	20.70	30.95	5.35	7.74
BRS when necessary	18	18	92.44	79.44	14.55	30.68	3.43	7.23

***Table 5.* T-test**

Table 5

Independent Samples T-test Results for Measures Related to BRS Decision

	95% CI					
	LB	UB	t	*df*	*p*	η^2
Looked for place to land	1.93	40.57	2.26	29	.02	.15
BRS when necessary	-3.26	29.26	1.63	34	.06	.07

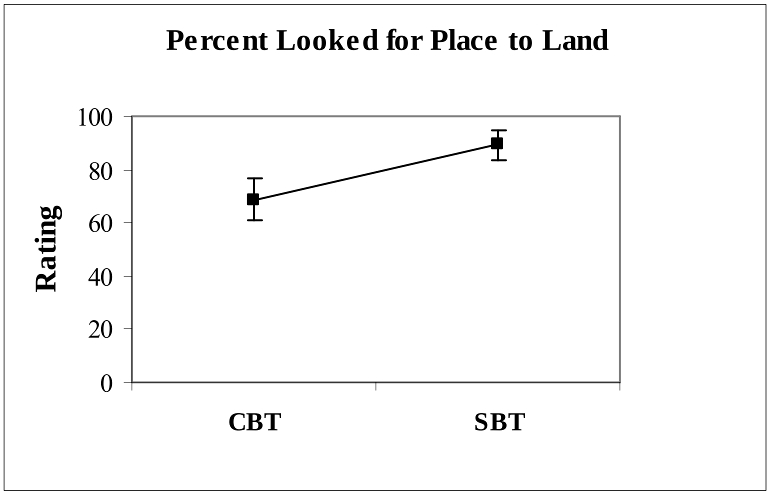


Figure 6. Mean and standard error of the mean for the performance measure “Looked for place to land”, with rating as a function of condition.

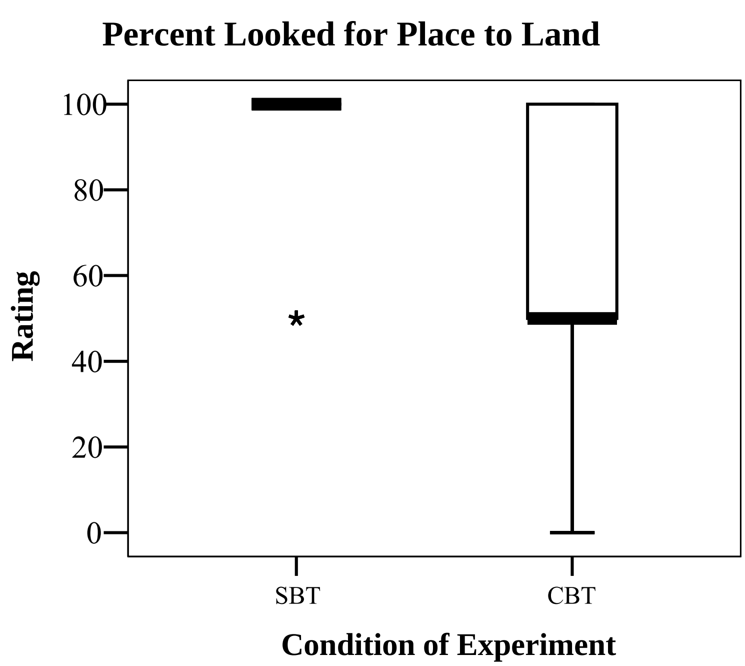


Figure 7. Box plot for the performance measure “Looked for place to land”, with rating as a function of condition.

Figure 8 depicts the posttest mean and standard error of the mean for the "Percent used BRS parachute when necessary" measure. Figure 9 contains a box plot for both conditions. This subhypothesis was analyzed with an independent samples t-test. As shown in Table 5, a significant difference was not found between groups, $t(36) = 1.63$, $p = .06$. Thus, condition was not significantly related to ratings of "BRS when necessary" (subhypothesis 1c). Therefore, the SBT group did not perform significantly better and this subhypothesis was not supported.

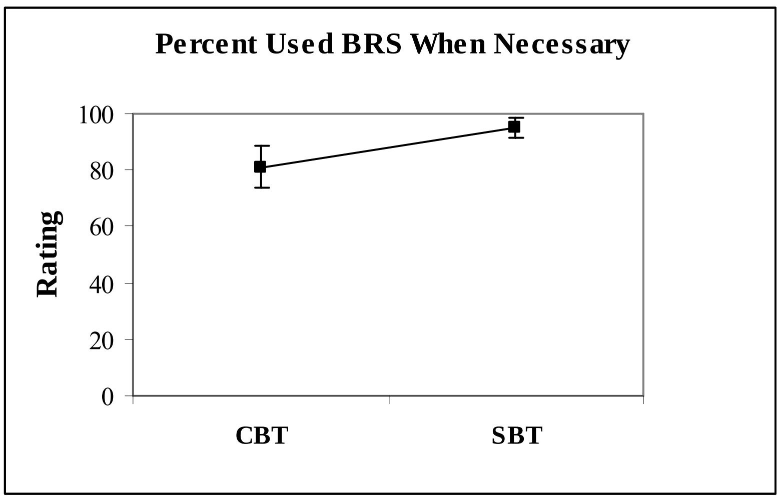


Figure 8. Mean and standard error of the mean for the performance measure "BRS when necessary", with rating as a function of condition.

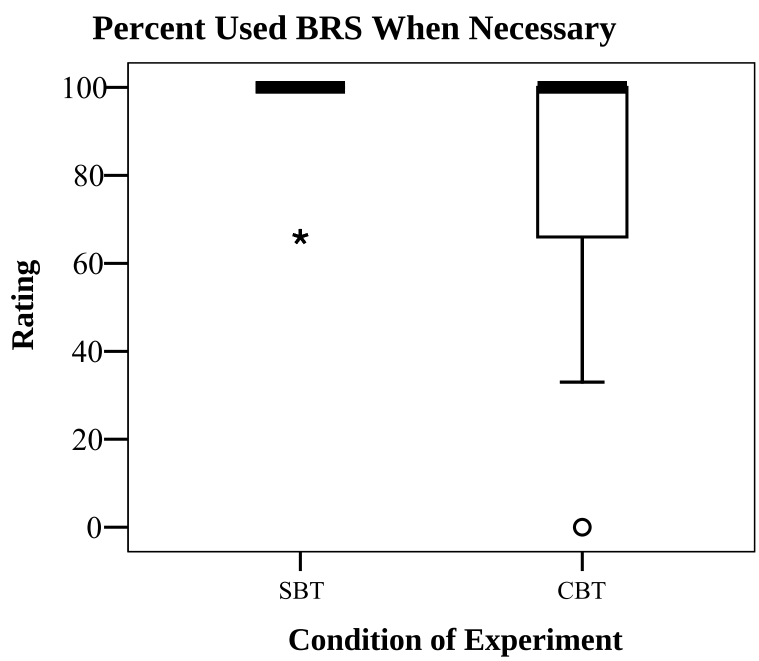


Figure 9. Box plot for the performance measure “BRS When Necessary”, with rating as a function of condition.

It was hypothesized that participants in the SBT condition would perform significantly better than participants in the CBT condition on the performance measure “Pilot uses parachute correctly”, also referred to as “correct BRS use” (subhypothesis 1d). The means and standard deviations for the pre and posttest performance measures for both conditions are shown in Table 2. Figure 10 depicts the mean and standard error of the mean for session and condition. Figures 11 and 12 contain box plots for the pre and posttest. This subhypothesis was analyzed with a two-way ANOVA. As shown in Table 6, a main effect was found for session, $f(1, 27) = 74.30$, $p = .00$. Thus, session was significantly related to ratings of “correct BRS use”. An eta squared of .73 indicates that 73% of the variability in “correct BRS use” is related to differences in session. The means of the posttest were higher than the means of the pretest across condition.

However, a significant effect was not found for condition, f (1, 27) = .21, p = .65. The interaction for this performance measure was also not significant, f (1, 27) = .07, p = .80. Thus, this subhypothesis was not supported.

***Table 6.* ANOVA**

Table 6

Two way Analysis of Variance for "Pilot Uses Parachute Correctly" Performance Measure

Source	*df*	F	η^2	Power	*p*
Session (S)	1	74.30	.73	1.00	.00
Group (G)	1	.21	.00	.07	.65
S × G	1	.07	.00	.06	.80
S within-group error	27	(1.27)			

Note. Values enclosed in parentheses represent mean square errors.

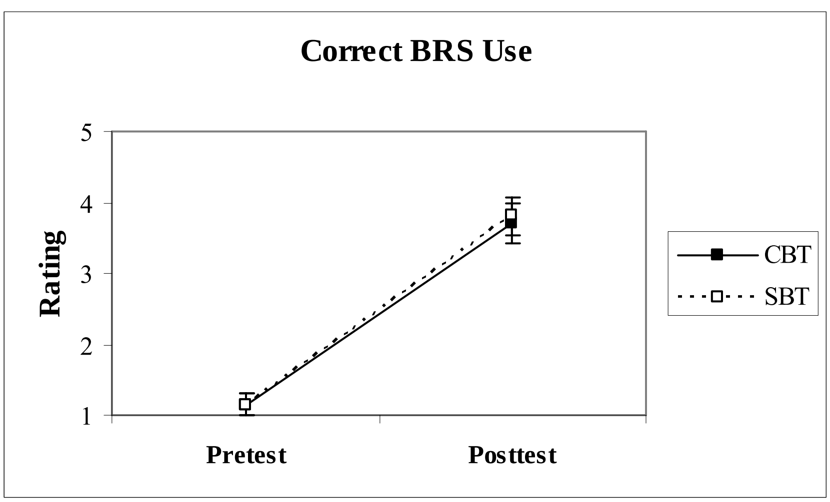


Figure 10. Mean and standard error of the mean for the performance measure "Correct BRS use", with rating as a function of session and condition.

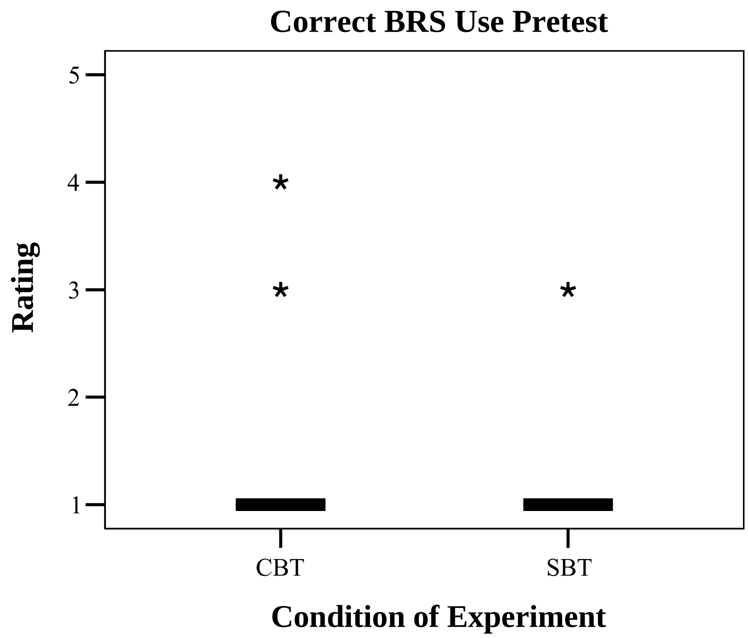


Figure 11. Box plot for the performance measure “Correct BRS use”, with pretest rating as a function of condition.

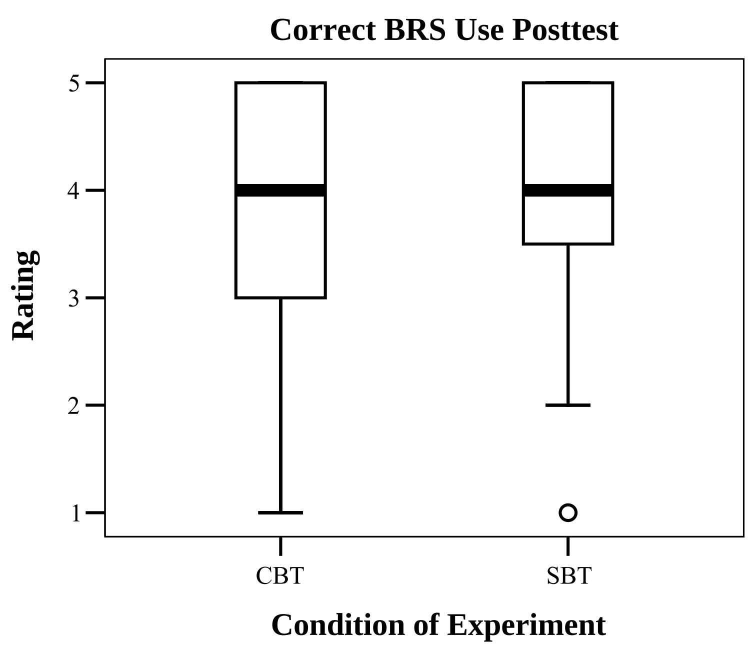


Figure 12. Box plot for the performance measure “Correct BRS use”, with posttest rating as a function of condition.

It was hypothesized that participants in the SBT condition would perform significantly better than participants in the CBT condition on the performance measure “Pilot utilizes 5 P/SRM technique appropriately” (subhypothesis 1e). The mean and standard deviation for the pre and posttest performance measures for both conditions is shown in Table 2. Figure 13 depicts the mean and standard error of the mean for session and condition. Figures 14 and 15 contain box plots for the pre and posttest. This subhypothesis was analyzed with a two-way ANOVA. As shown in Table 7, a main effect was found for session, $f(1, 31) = 61.55$, $p = .00$. Thus, session was significantly related to ratings of “SRM”. An eta squared of .32 indicates that 32% of the variability in “SRM” is related to differences in session. The means of the posttest were higher than the means of the pretest across condition. A significant main effect was also found for condition, $f(1, 31) = 50.60$, $p = .00$. Thus, condition was significantly related to ratings of “SRM”. An eta squared of .26 indicates that 26% of the variability in “SRM” is related to differences in condition. The SBT group had an overall higher mean across time (pre-post). The interaction for this performance measure was also significant, $f(1, 31) = 50.60$, $p = .00$. Thus, the interaction was significantly related to ratings of “SRM”. An eta squared of .26 indicates that 26% of the variability in “SRM” is related to differences caused by the interaction (i.e. the combined effect of training method and session). The Tukey/Kramer method of comparison (.05) was used for further analysis of the interaction, as it is the preferred method when uneven n’s are present. First, the pretest means for SBT and CBT groups were exactly the same ($X = 1$), so a significant difference was not found in the pairwise comparison. The SBT posttest mean was again significantly higher than the SBT pretest mean ($p = .00$). The means for the CBT pre and posttest were not significantly different (p = .60). When comparing the SBT posttest mean to the CBT posttest mean, the SBT mean was higher than the CBT mean, showing the SBT group received significantly higher performance ratings for SRM (p = .00). Therefore, participants in the SBT condition performed significantly better than participants in the CBT condition on this performance measure and this subhypothesis was supported.

***Table 7.* ANOVA**

Table 7

Two way Analysis of Variance for "Pilot Utilizes 5 P's/SRM Appropriately" Performance Measure

Source	*df*	F	η^2	Power	*p*
Session (S)	1	61.55	.32	1.00	.00
Group (G)	1	50.60	.26	1.00	.00
S × G	1	50.60	.26	1.00	.00
S within-group error	31	(.35)			

Note. Values enclosed in parentheses represent mean square errors.

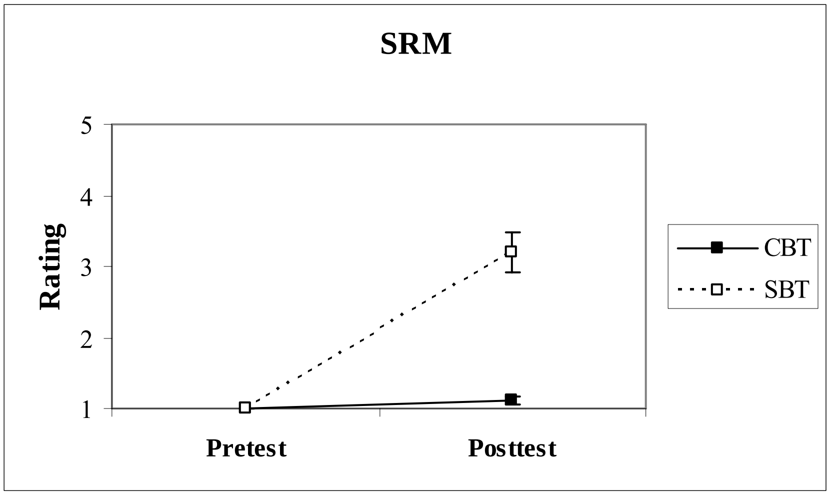


Figure 13. Mean and standard error of the mean for the performance measure “SRM”, with rating as a function of session and condition.

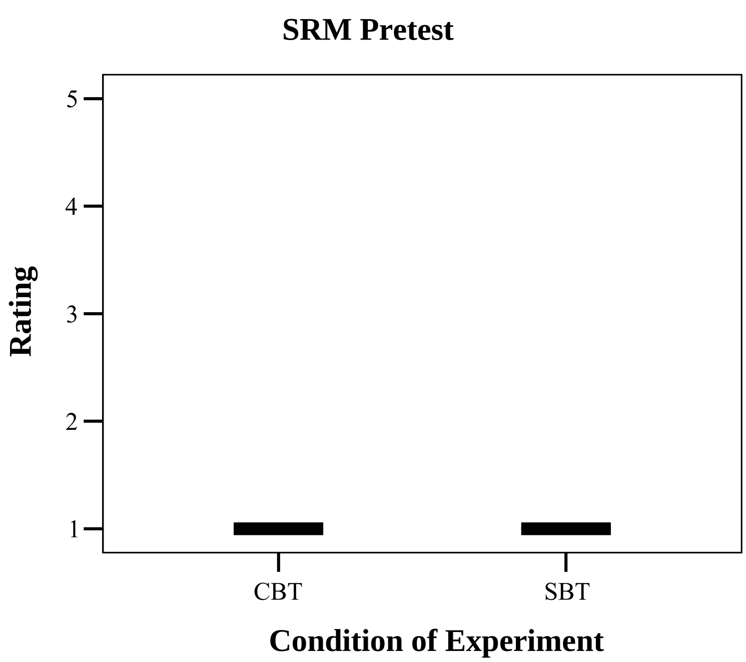


Figure 14. Box plot for the performance measure “SRM”, with pretest rating as a function of condition.

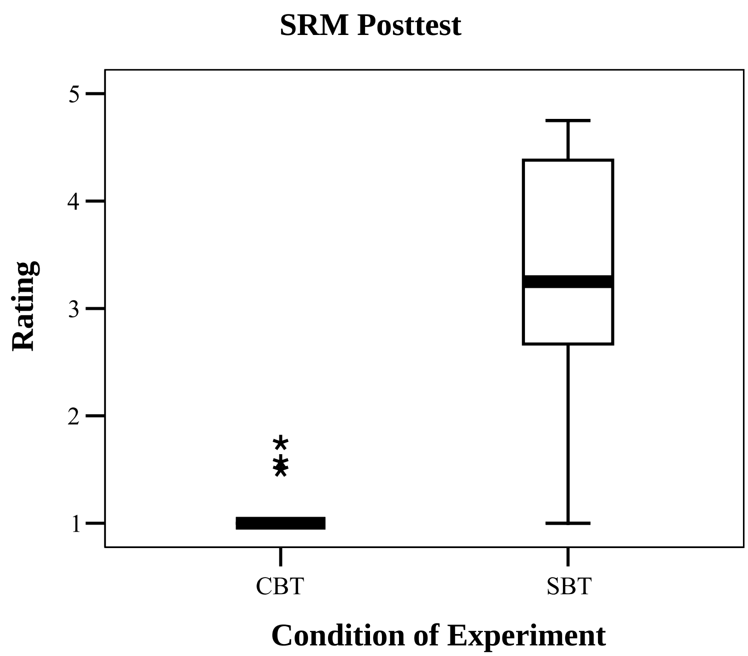


Figure 15. Box plot for the performance measure "SRM", with posttest rating as a function of condition.

It was hypothesized that participants in the SBT condition would perform significantly better than participants in the CBT condition on the performance measure "Overall the pilot responded" (subhypothesis 1f). The mean and standard deviation for the pre and posttest performance measures for both conditions is shown in Table 2. Figure 16 depicts the mean and standard error of the mean for session and condition. Figures 17 and 18 contain box plots for the pre and posttest. This subhypothesis was analyzed with a two-way ANOVA. As shown in Table 8, a main effect was found for session, $f(1, 32) = 17.32$, $p = .00$. Thus, session was significantly related to ratings of "overall performance". An eta squared of .29 indicates that 29% of the variability in "overall performance" is related to differences in session. The means of the posttest were higher than the means of the pretest across condition. A significant main effect was not found for condition, $f(1, 32) = 3.85$, $p = .06$. However, the interaction for this performance measure was significant, $f(1, 32) = 6.72$, $p = .01$. Thus, the ratings of overall performance were significantly impacted by the interaction (i.e. the combined effect of training method and session). An eta squared of .11 indicates that 11% of the variability in “overall performance” is related to differences caused by interaction. The Tukey HSD (honestly significant difference) method of comparison (.05) was used for the post hoc analysis. First, no significant difference was found in the pairwise comparison of pretest means for SBT and CBT, indicating the groups were equivalent before the training ($p = .72$). The SBT group had significantly different pre and posttest means ($p = .00$). As shown in Table 2, the mean for the SBT posttest was higher than the SBT pretest. The CBT group did not have significantly different pre and posttest means ($p = .28$). Significance was also found when comparing the SBT posttest mean to the CBT posttest mean ($p = .00$). The SBT mean was higher than the CBT mean, showing the SBT group received significantly higher performance ratings for SRM. Therefore, SBT performed significantly better than the CBT group on this performance measure and this subhypothesis was supported.

***Table 8.* ANOVA**

Table 8

Two way Analysis of Variance for "Overall the Pilot Responded" Performance Measure

Source	*Df*	F	η^2	Power	*p*
Session (S)	1	17.32	.29	.98	.00
Group (G)	1	3.85	.06	.48	.06
S × G	1	6.72	.11	.71	.01
S within-group error	32	(.60)			

Note. Values enclosed in parentheses represent mean square errors.

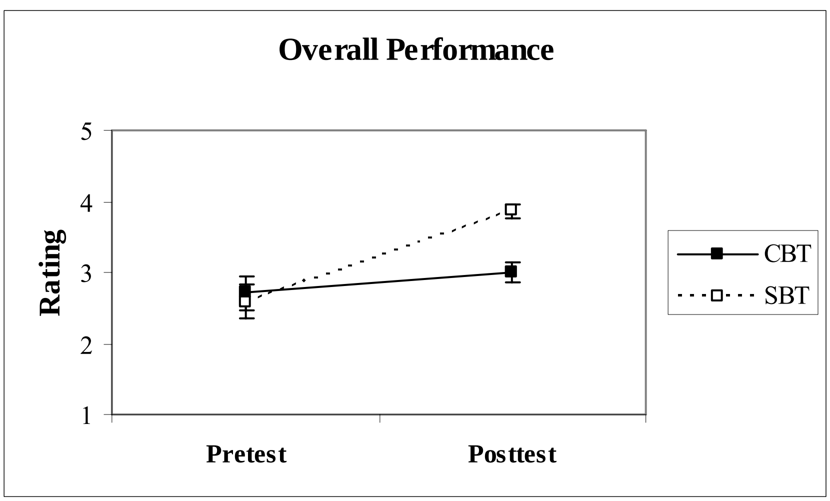


Figure 16. Mean and standard error of the mean for the performance measure "overall performance", with rating as a function of session and condition.

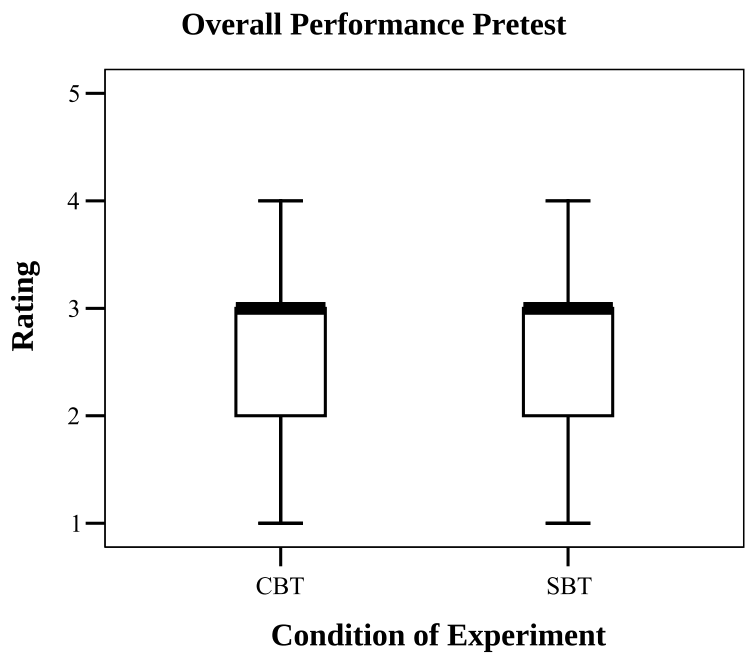


Figure 17. Box plot for the performance measure "overall performance", with pretest rating as a function of condition.

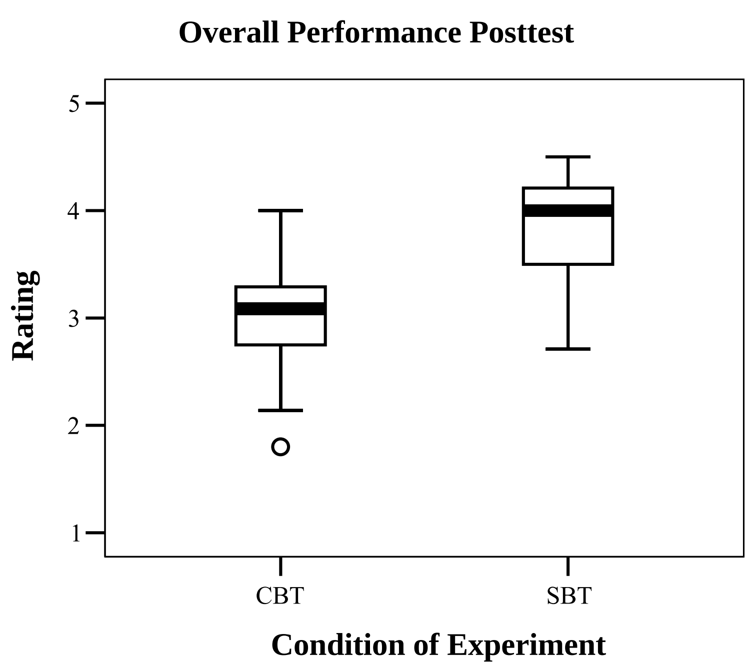


Figure 18. Box plot for the performance measure "overall performance", with posttest rating as a function of condition.

It was hypothesized that participants in the SBT condition would perform significantly better than participants in the CBT condition on the performance measure “Frequency did not crash” (subhypothesis 1g). A percentage was calculated for this measure based on the frequency participants did *not* crash. As with the other measures, on this scale a high score is desirable. The mean and standard deviation for the pre and posttest performance measures for both conditions is shown in Table 2. Figure 19 depicts the mean and standard error of the mean for session and condition. Figures 20 and 21 contain box plots for the pre and posttest. This subhypothesis was analyzed with a two-way ANOVA. As shown in Table 9, a main effect was found for session, $f(1, 32) = 7.01$, $p = .01$. Thus, session was significantly related to ratings of “frequency did not crash”. An eta squared of .17 indicates that 17% of the variability in “frequency did not crash” is related to differences in session. The means of the posttest were higher than the means of the pretest across condition. A significant main effect was not found for condition, $f(1, 32) = 1.81$, $p = .19$. The interaction for this performance measure was also not significant, $f(1, 32) = .00$, $p = .99$. Thus, this subhypothesis was not supported.

***Table 9.* ANOVA**

Table 9

Two way Analysis of Variance for “Frequency Did Not Crash” Performance Measure

Source	*df*	F	η^2	Power	*p*
Session (S)	1	7.01	.17	.73	.01
Group (G)	1	1.81	.04	.26	.19
S × G	1	.00	.00	.05	.99
S within-group error	32	(185.86)			

Note. Values enclosed in parentheses represent mean square errors.

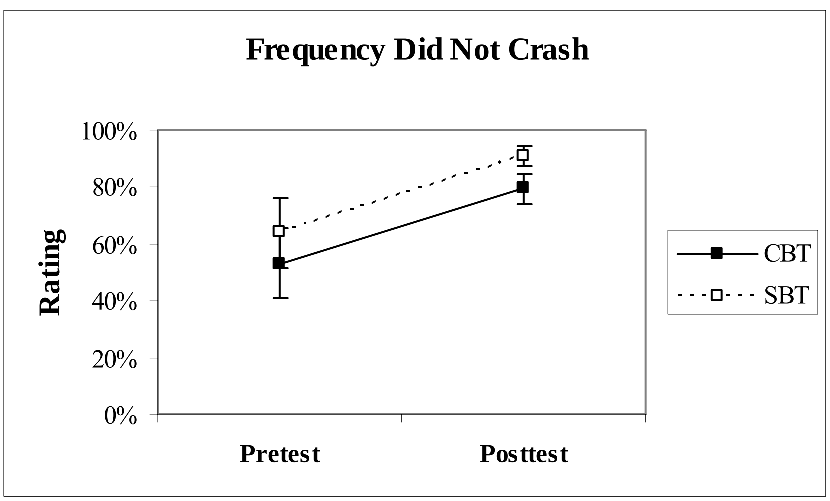


Figure 19. Mean and standard error of the mean for the performance measure “frequency did not crash”, with rating as a function of session and condition.

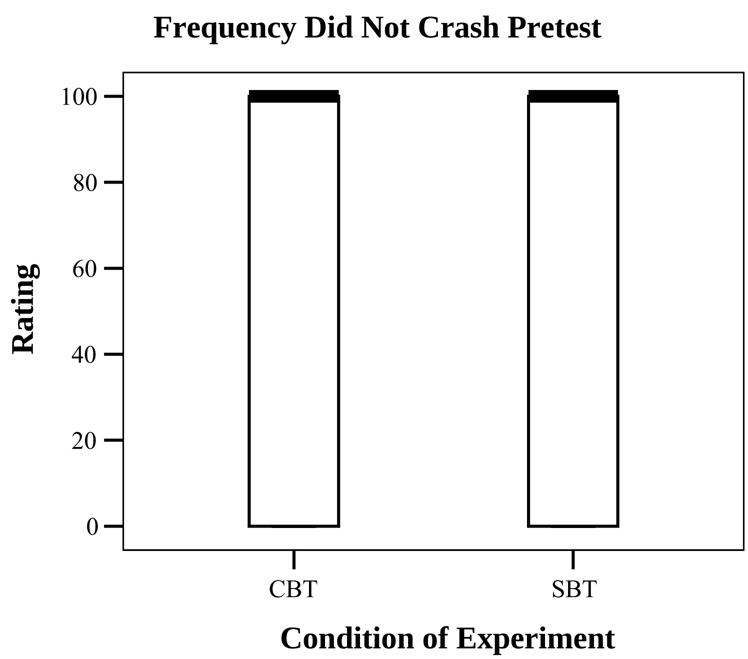


Figure 20. Box plot for the performance measure “frequency did not crash”, with pretest rating as a function of condition.

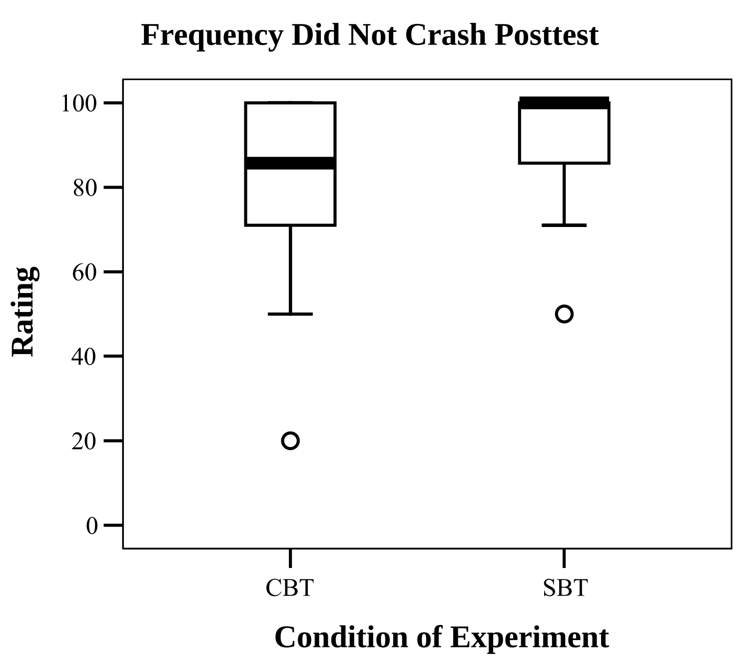


Figure 21. Box plot for the performance measure “frequency did not crash”, with posttest rating as a function of condition.

It was hypothesized that participants in the SBT condition would perform significantly better than participants in the CBT condition on the performance measure "Pilot uses parachute at correct time based on the sequence of events" (subhypothesis 1h). This performance measure will be referred to as "BRS timing" to abbreviate the title. Pretest data was not available for this performance measure. However, the mean and standard deviation for the posttest data is shown for both groups in Table 10. Figure 22 depicts the posttest mean and standard error of the mean for both conditions. Figure 23 contains a box plot for both conditions. This subhypothesis was analyzed with an independent samples t-test. As shown in Table 11, a significant difference was found between groups, $t\ (33) = 3.65$, $p = .00$. Thus, condition was significantly related to ratings of "BRS timing". An eta squared of .10 indicates that 10% of the variability in "BRS timing" is related to differences in condition. The mean of the SBT condition was higher than the mean of the CBT condition. Therefore, the SBT group performed significantly better and this subhypothesis was supported.

***Table 10.* Descriptive Statistics**

Table 10

Means and Standard Deviations of Performance Data Posttest Ratings on Performance Measures Without Pretest Data

Performance Measure	*n*		*M*		*SD*		*St Error Mean*	
	SBT	CBT	SBT	CBT	SBT	CBT	SBT	CBT
BRS timing	17	18	4.27	2.85	.75	1.43	.18	.34
BRS altitude	17	17	4.47	3.73	1.05	1.38	.25	.33
BRS knots	16	17	3.94	4.65	1.69	1.00	.42	.24
Control plane	18	18	4.97	4.78	.12	.35	.03	.08
Diverts	18	18	3.97	3.06	1.05	1.24	.25	.29
Contact ATC	18	18	3.76	2.66	.98	1.18	.23	.28
Declare emergency	18	18	4.00	2.33	1.32	1.19	.31	.28
Checklist	18	18	2.35	1.80	1.28	1.01	.30	.24
Checklist steps	18	18	2.05	1.45	.92	.77	.22	.18

***Table 11.* T-test**

Table 11

Independent Samples T-test Results for Comparison of SBT and CBT Posttest Means for Performance Measures Without Corresponding Pretest Data

Performance Measure	95% CI					
	LB	UB	T	*df*	*p*	η^2
BRS timing	.63	2.22	3.65	33	.00	.10
BRS altitude	-.11	1.60	1.78	32	.04	.05
BRS knots	-1.69	.27	-1.48	31	.08	.05
Control plane	.02	.37	2.22	34	.02	.06
Diverts	.14	1.69	2.40	34	.01	.07
Contact ATC	.37	1.84	3.05	34	.00	.08
Declare emergency	.82	2.52	3.99	34	.00	.11
Checklist	-.23	1.34	1.44	34	.08	.05
Checklist steps	.03	1.18	2.13	34	.04	.06

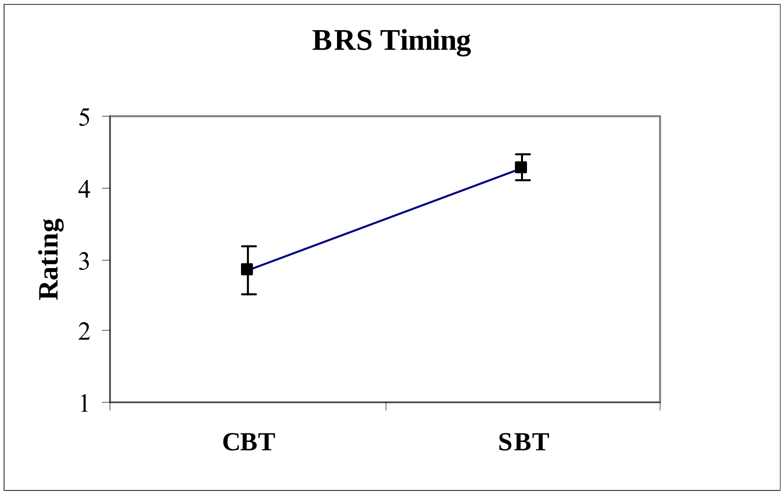


Figure 22. Mean and standard error of the mean for the posttest performance measure "BRS timing", with rating as a function of condition.

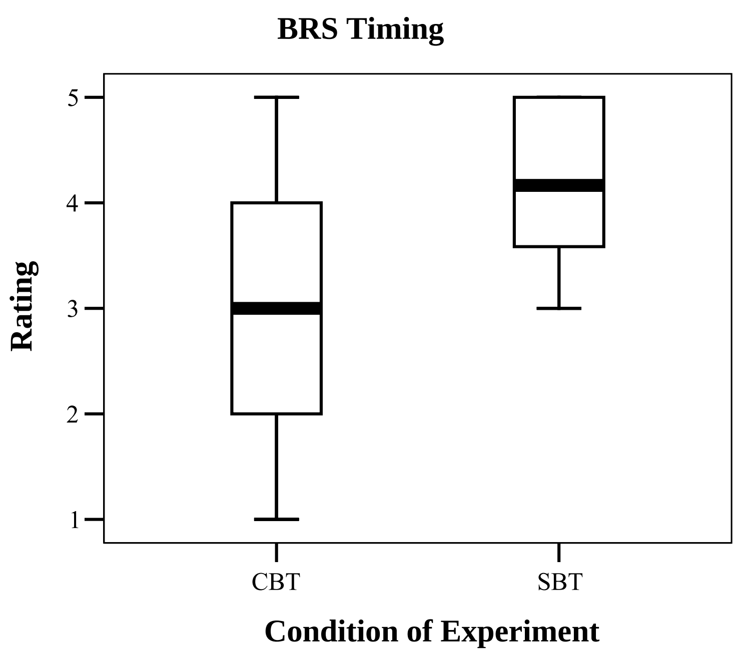


Figure 23. Box plot for the posttest performance measure "BRS timing", with rating as a function of condition.

It was hypothesized that participants in the SBT condition would perform significantly better than participants in the CBT condition on the performance measure "Pilot uses parachute above the minimum altitude (500 ft)" (subhypothesis 1i), which will be referred to as simply "BRS altitude". Pretest data was not available for this performance measure. However, the mean and standard deviation for the posttest data is shown for both groups in Table 10. Figure 24 depicts the posttest mean and standard error of the mean for both conditions. Figure 25 contains a box plot for both conditions. This subhypothesis was analyzed with an independent samples t-test. As shown in Table 11, a significant difference was found between groups, t (32) = 1.78, p = .04. Thus, condition was significantly related to ratings of "BRS altitude". An eta squared of .05 indicates that 5% of the variability in "BRS altitude" is related to differences in condition. The mean of the SBT condition was higher than the mean of the CBT condition. Therefore, the SBT group performed significantly better and this subhypothesis was supported.

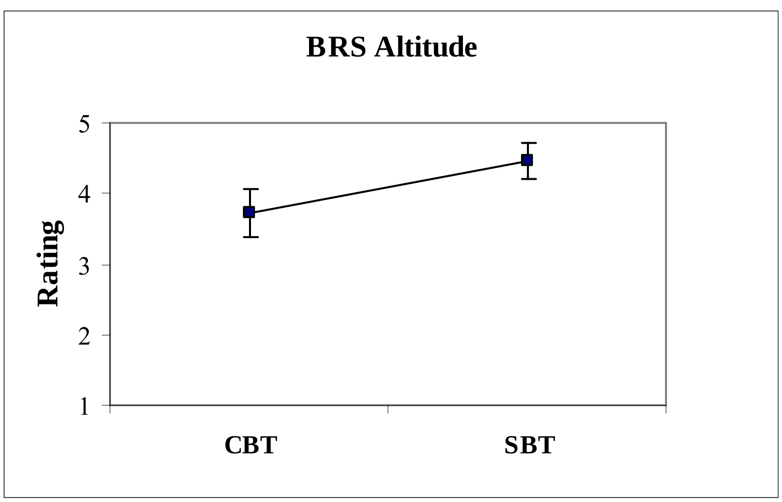


Figure 24. Mean and standard error of the mean for the posttest performance measure "BRS altitude", with rating as a function of condition.

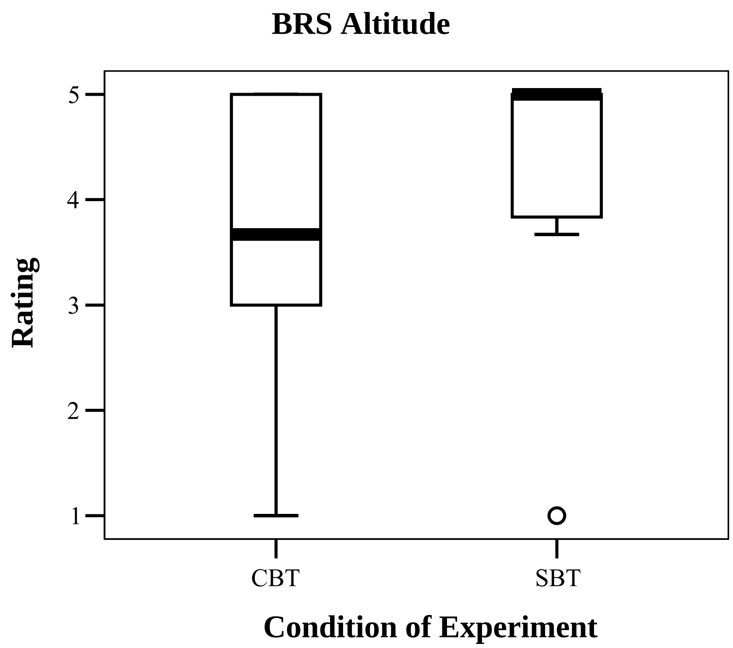


Figure 25. Box plot for the posttest performance measure “BRS altitude”, with rating as a function of condition.

It was hypothesized that participants in the SBT condition would perform significantly better than participants in the CBT condition on the performance measure “Pilot uses parachute below the maximum rate (90 knots)”, also referred to as “BRS knots” (subhypothesis 1j). The maximum rate is actually 135 knots. However, an artificial limit of 90 knots was established for this study because the typical cruise rate for a Cessna 172 is well below 135 knots. Pretest data was not available for this performance measure. However, the mean and standard deviation for the posttest data is shown for both groups in Table 10. Figure 26 depicts the posttest mean and standard error of the mean for both conditions. Figure 27 contains a box plot for both conditions. This subhypothesis was analyzed with an independent samples t-test. As shown in Table 11, a significant difference was not found between groups, $t(31) = -1.48$, $p = .08$.

Therefore, the SBT group did not perform significantly better and this subhypothesis was not supported.

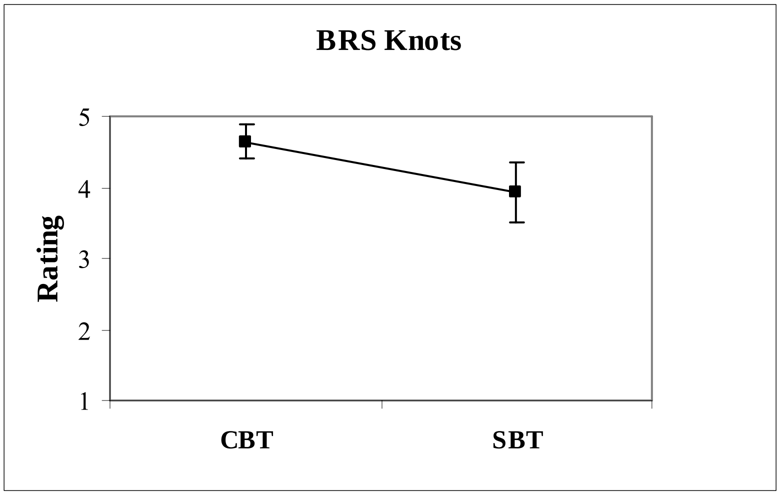


Figure 26. Mean and standard error of the mean for the posttest performance measure "BRS knots", with rating as a function of condition.

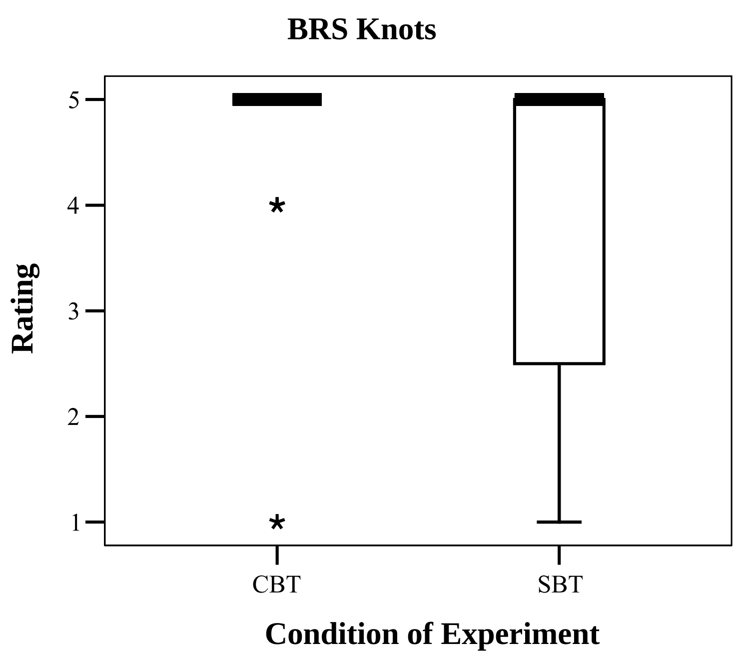


Figure 27. Box plot for the posttest performance measure "BRS knots", with rating as a function of condition.

In order to measure performance in greater detail, some emergency responses were measured that did not involve the BRS parachute or SRM. Data was also collected on the following performance measures, although they are not directly related to parachute use:

It was hypothesized that participants in the SBT condition would perform significantly better than participants in the CBT condition on the performance measure “Pilot maintains control of the aircraft” (subhypothesis 1k). Pretest data was not available for this performance measure. However, the mean and standard deviation for the posttest data is shown for both groups in Table 10. Figure 28 depicts the posttest mean and standard error of the mean for both conditions. Figure 29 contains a box plot for both conditions. This subhypothesis was analyzed with an independent samples t-test. As shown in Table 11, a significant difference was found between groups, $t\ (34) = 2.22$, $p = .02$. Thus, condition was significantly related to ratings of “control aircraft”. An eta squared of .06 indicates that 6% of the variability in “control aircraft” is related to differences in condition. The mean of the SBT condition was higher than the mean of the CBT condition. Thus, the SBT group performed significantly better and this subhypothesis was supported.

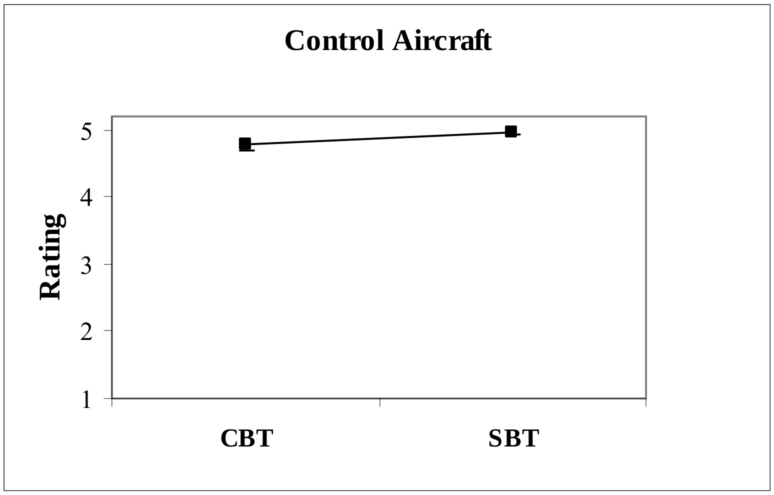


Figure 28. Mean and standard error of the mean for the posttest performance measure "control aircraft", with rating as a function of condition.

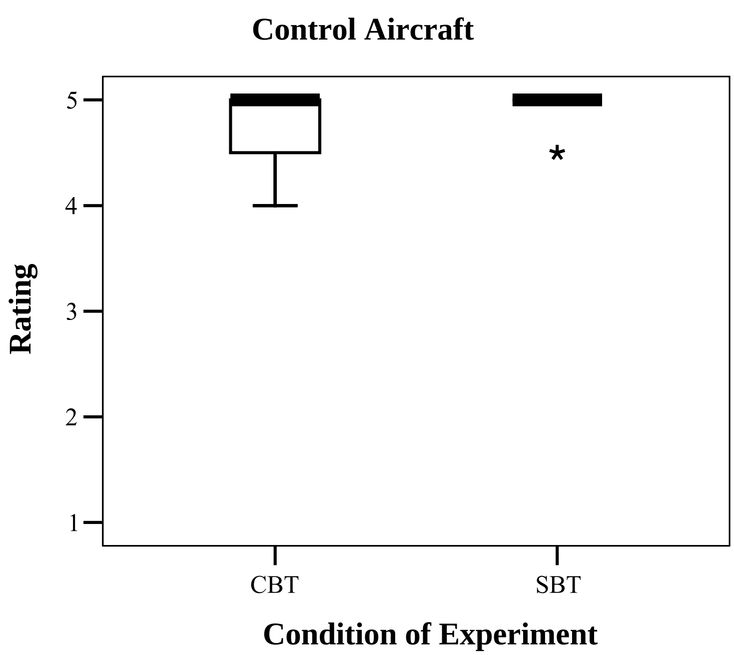


Figure 29. Box plot for the posttest performance measure "control aircraft", with rating as a function of condition.

It was hypothesized that participants in the SBT condition would perform significantly better than participants in the CBT condition on the performance measure "Pilot refers to checklist to resolve problem" (subhypothesis 1l). Pretest data was not available for this performance measure. However, the mean and standard deviation for the posttest data is shown for both groups in Table 10. Figure 30 depicts the posttest mean and standard error of the mean for both conditions. Figure 31 contains a box plot for both conditions. This subhypothesis was analyzed with an independent samples t-test. As shown in Table 11, a significant difference was not found between groups, t (34) = 1.44, p = .08. Therefore, the SBT group did not perform significantly better and this subhypothesis was not supported.

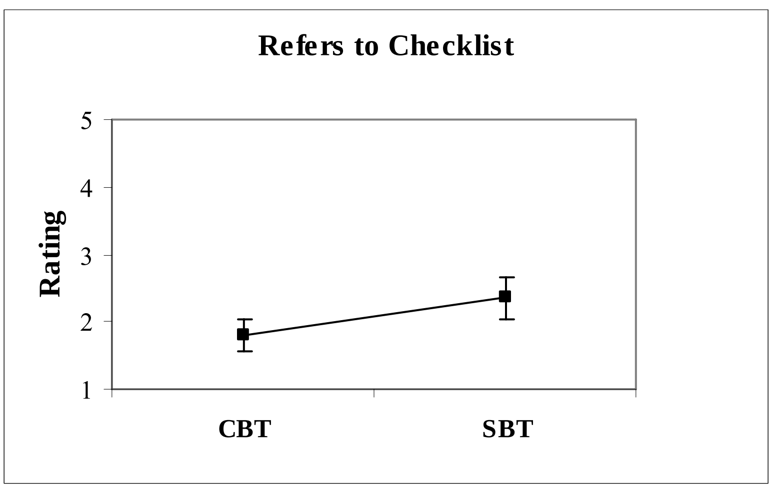


Figure 30. Mean and standard error of the mean for the posttest performance measure "refers to checklist", with rating as a function of condition.

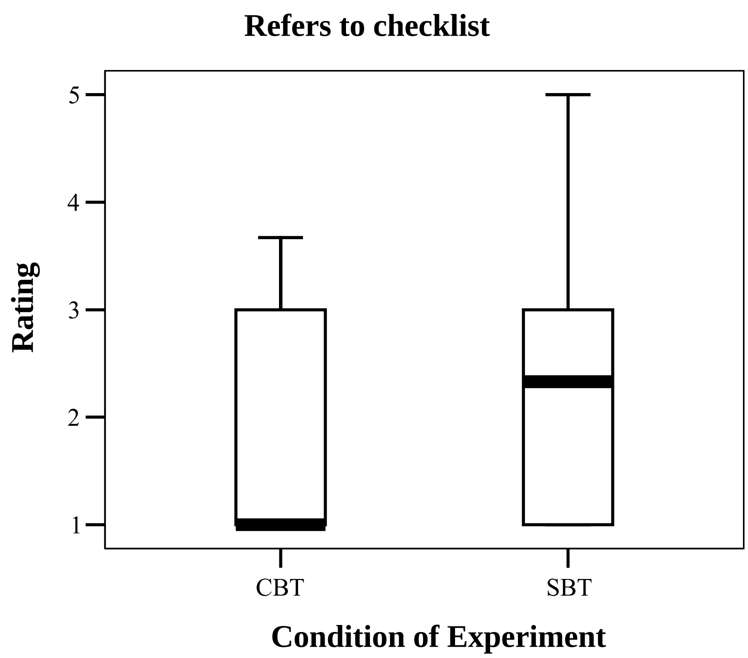


Figure 31. Box plot for the posttest performance measure “refers to checklist”, with rating as a function of condition.

It was hypothesized that participants in the SBT condition would perform significantly better than participants in the CBT condition on the performance measure “Pilot follows checklist procedure to resolve problem” (subhypothesis 1m). Pretest data was not available for this performance measure. However, the mean and standard deviation for the posttest data is shown for both groups in Table 10. Figure 32 depicts the posttest mean and standard error of the mean for both conditions. Figure 33 contains a box plot for both conditions. This subhypothesis was analyzed with an independent samples t-test. As shown in Table 11, a significant difference was found between groups, $t\ (34) = 2.13$, $p = .02$. Thus, condition was significantly related to ratings of “follows checklist”. An eta squared of .06 indicates that 6% of the variability in “follows checklist” is related to differences in condition. The mean of the SBT condition was higher than the mean of the CBT condition. Therefore, the SBT group performed significantly better and this subhypothesis was supported.

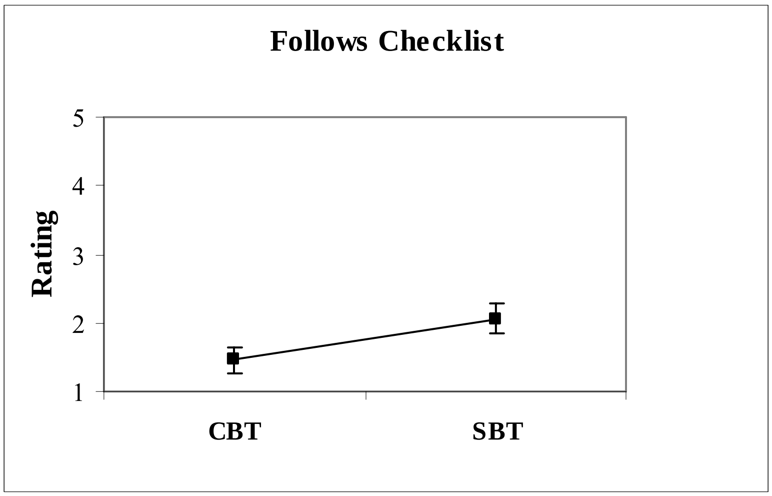


Figure 32. Mean and standard error of the mean for the posttest performance measure “follows checklist”, with rating as a function of condition.

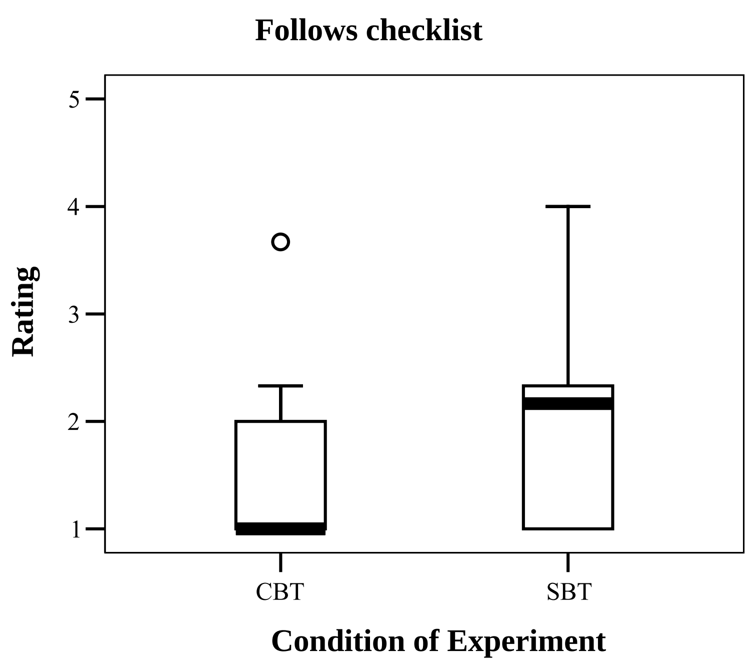


Figure 33. Box plot for the posttest performance measure “follows checklist”, with rating as a function of condition.

It was hypothesized that participants in the SBT condition would perform significantly better than participants in the CBT condition on the performance measure "Pilot contacts ATC" (subhypothesis 1n). Pretest data was not available for this performance measure. However, the mean and standard deviation for the posttest data is shown for both groups in Table 10. Figure 34 depicts the posttest mean and standard error of the mean for both conditions. Figure 35 contains a box plot for both conditions. This subhypothesis was analyzed with an independent samples t-test. As shown in Table 11, a significant difference was found between groups, $t(34) = 3.05$, $p = .00$. Thus, condition was significantly related to ratings of "contacts ATC". An eta squared of .08 indicates that 8% of the variability in "contacts ATC" is related to differences in condition. The mean of the SBT condition was higher than the mean of the CBT condition. Therefore, the SBT group performed significantly better and this subhypothesis was supported.

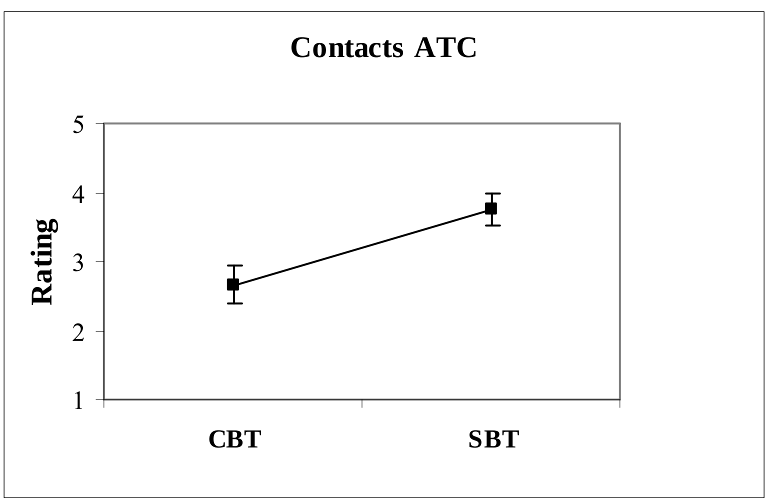


Figure 34. Mean and standard error of the mean for the posttest performance measure "contacts ATC", with rating as a function of condition.

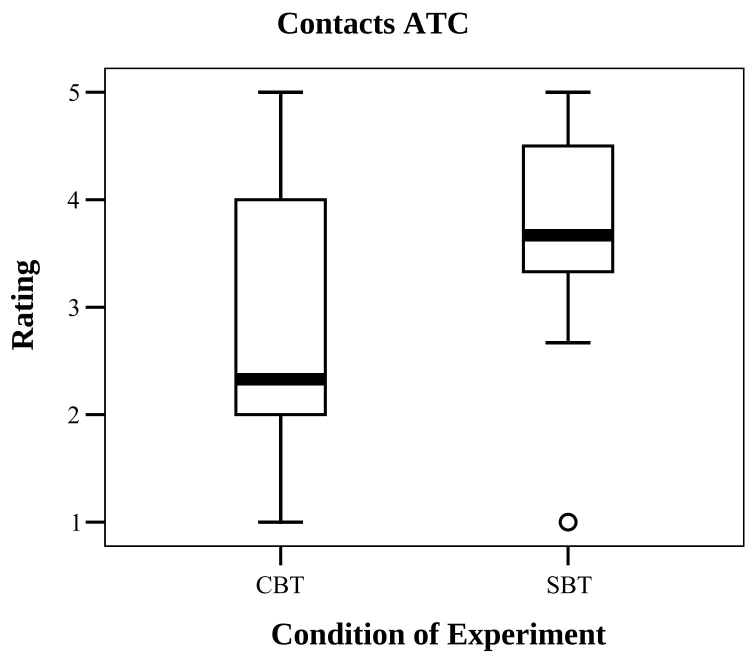


Figure 35. Box plot for the posttest performance measure "contacts ATC", with rating as a function of condition.

It was hypothesized that participants in the SBT condition would perform significantly better than participants in the CBT condition on the performance measure "Pilot declares an emergency" (subhypothesis 1o). Pretest data was not available for this performance measure. However, the mean and standard deviation for the posttest data is shown for both groups in Table 10. Figure 36 depicts the posttest mean and standard error of the mean for both conditions. Figure 37 contains a box plot for both conditions. This subhypothesis was analyzed with an independent samples t-test. As shown in Table 11, a significant difference was found between groups, t (34) = 3.99, p = .00. Thus, condition was significantly related to ratings of "declares emergency". An eta squared of .11 indicates that 11% of the variability in "declares emergency" is related to differences in condition. The mean of the SBT condition was higher than the mean of the CBT condition. Therefore, the SBT group performed significantly better and this subhypothesis was supported.

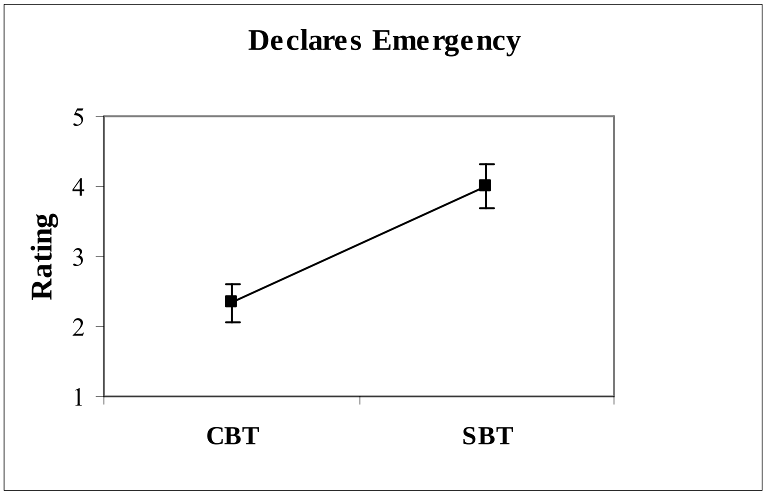


Figure 36. Mean and standard error of the mean for the posttest performance measure "declares emergency", with rating as a function of condition.

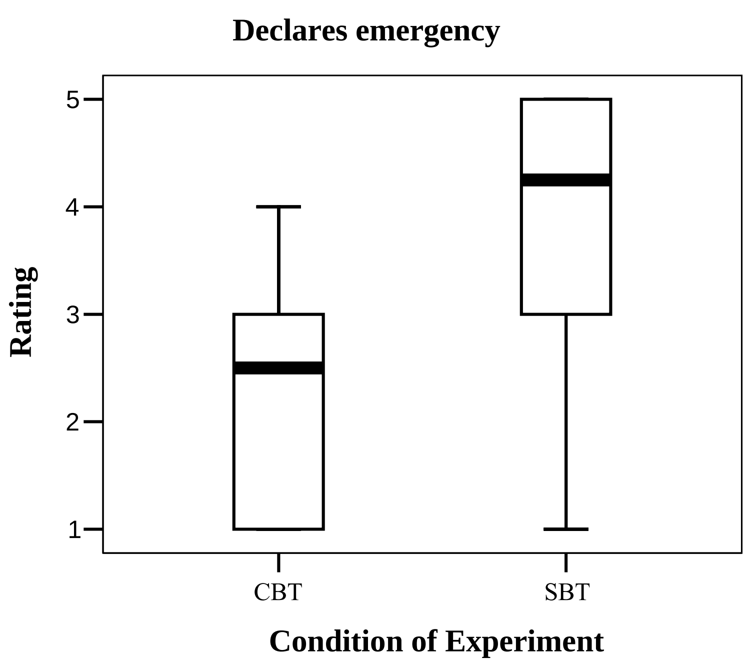


Figure 37. Box plot for the posttest performance measure "declares emergency", with rating as a function of condition.

It was hypothesized that participants in the SBT condition would perform significantly better than participants in the CBT condition on the performance measure "Pilot diverts or continues to destination around storm" (subhypothesis 1p). Pretest data was not available for this performance measure. However, the mean and standard deviation for the posttest data is shown for both groups in Table 10. Figure 38 depicts the posttest mean and standard error of the mean for both conditions. Figure 39 contains a box plot for both conditions. This subhypothesis was analyzed with an independent samples t-test. As shown in Table 11, a significant difference was found between groups, t (34) = 2.40, p = .01. Thus, condition was significantly related to ratings of "diverts". An eta squared of .07 indicates that 7% of the variability in "diverts" is related to differences in condition. The mean of the SBT condition was higher than the mean of the CBT condition. Therefore, the SBT group performed significantly better and this subhypothesis was supported.

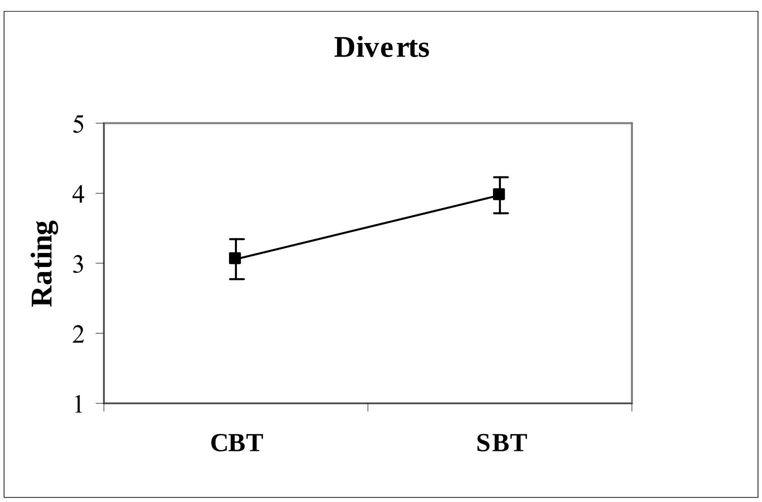


Figure 38. Mean and standard error of the mean for the posttest performance measure "diverts", with rating as a function of condition.

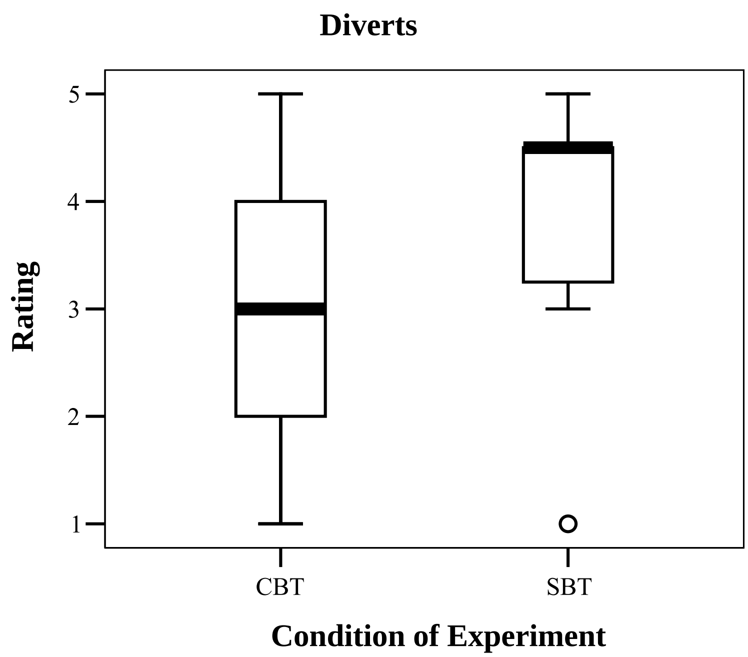


Figure 39. Box plot for the posttest performance measure “diverts”, with rating as a function of condition.

In summary, no significant differences were found in the comparison of the SBT and CBT pretest measures, indicating the two groups had equivalent performance levels prior to being trained. Overall, the SBT condition performed significantly better than the CBT condition on ten of the 16 subhypotheses. Thus, hypothesis 1 was generally supported.

Knowledge Test

Hypothesis 2: Participants in the SBT condition will achieve higher scores on the knowledge test than the participants in the control condition. To test this hypothesis, mean scores between groups on a knowledge test were compared using an independent samples t-test. The ten question knowledge test (given to participants at the end of the experiment) contained seven BRS and three SRM related questions. Scores for the seven

BRS related items were averaged and compared between groups. Scores for the three SRM related items were likewise averaged separately and also compared between groups. Means and standard deviations for both the BRS and SRM averaged test scores are reported in Table 12. Figure 40 depicts the posttest mean and standard error of the mean for the BRS portion of the knowledge test. Figure 41 contains a box plot for the BRS portion of the knowledge test. For the SRM portion of the knowledge test, Figure 42 depicts the posttest mean and standard error of the mean, while Figure 43 contains a box plot for both conditions.

Subhypothesis 2a: Table 13 contains the results of the independent samples t-tests performed on the averaged BRS and SRM questions. No significant differences were found between groups for the BRS portion of the knowledge test, $t\ (34) = -.69$, $p = .25$. Therefore, the scores for the SBT and the CBT groups did not differ in the BRS portion of the knowledge test and subhypothesis 2a was not supported.

Subhypothesis 2b: The averaged score for the SRM portion of the knowledge test was significantly different between groups, $t\ (34) = 6.29$, $p = .00$. Thus, condition was significantly related to SRM score on the knowledge test. An eta squared of .08 indicates that 8% of the variability in the SRM portion of the knowledge test is related to differences in condition. Specifically, the SBT condition scored higher on average than the CBT condition for SRM related questions. Thus, the SBT group outscored the CBT group on the SRM portion of the knowledge test and subhypothesis 2b was supported.

***Table 12.* Means and Standard Deviations**

Table 12

Means and Standard Deviations for BRS and SRM Averaged Knowledge Test Scores

Performance Measure	*n*		*M*		*SD*		*St Error Mean*	
	SBT	CBT	SBT	CBT	SBT	CBT	SBT	CBT
BRS average	18	18	17.44	18.00	2.77	2.00	.65	.47
SRM average	18	18	9.11	5.33	1.84	4.91	.43	1.15

***Table 13.* T-test**

Table 13

Independent Samples T-test Results for Comparison of SBT and CBT Means for BRS and SRM Knowledge Test

	95% CI					
	LB	UB	t	*df*	*p*	η^2
BRS averaged score	-2.19	1.08	-.69	34	.25	-.02
SRM averaged score	1.27	6.29	3.06	34	.00	.08

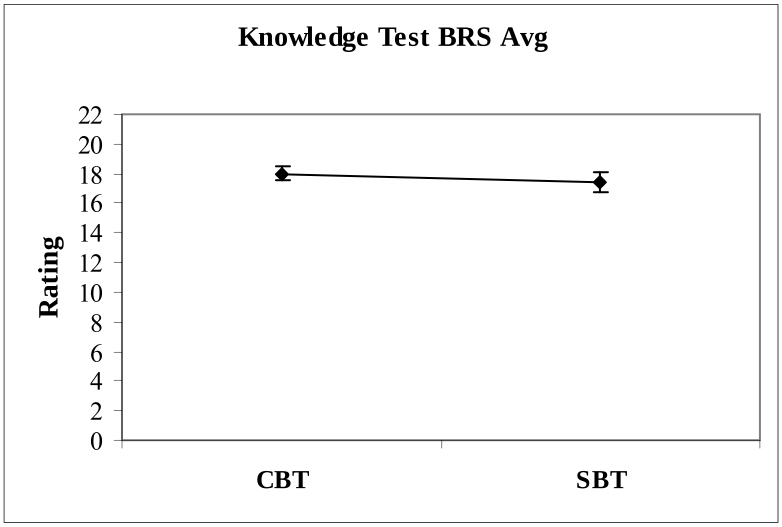


Figure 40. Mean and standard error of the mean for the BRS portion of the knowledge test, with rating as a function of condition.

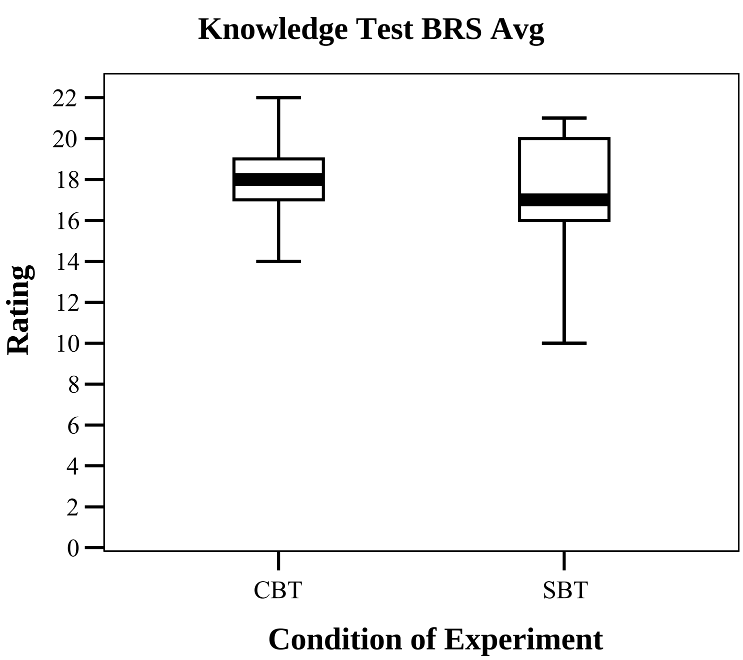


Figure 41. Box plot for the BRS portion of the knowledge test, with rating as a function of condition.

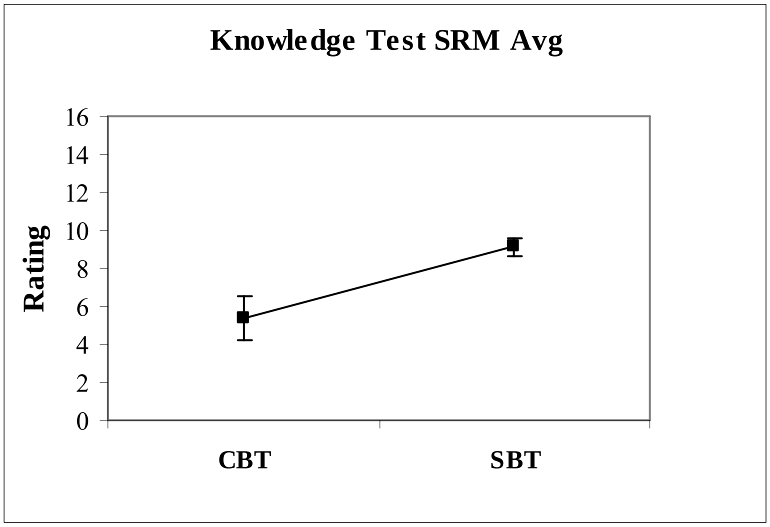


Figure 42. Mean and standard error of the mean for the SRM portion of the knowledge test, with rating as a function of condition.

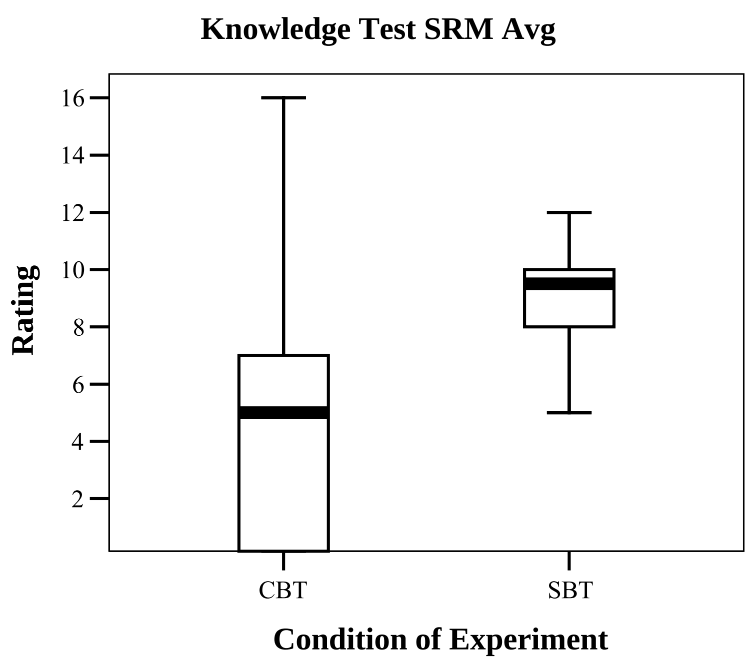


Figure 43. Box plot for the SRM portion of the knowledge test, with rating as a function of condition.

Self-Efficacy

Hypothesis 3: Participants in the SBT condition will have significantly higher self -efficacy regarding the use of the parachute and SRM than will the participants in the control condition. To test this hypothesis, mean ratings between groups for self-efficacy were compared using an independent samples t-test. The ten-item self-efficacy questionnaire contained seven BRS and three SRM related items. Therefore, ratings for the BRS and SRM items were averaged separately and compared between groups. Means and standard deviations for both the BRS and SRM averaged self-efficacy ratings are reported in Table 14. Figure 44 depicts the posttest mean and standard error of the mean for the BRS portion of the self-efficacy questionnaire. Figure 45 contains a box plot for the BRS portion of the self-efficacy questionnaire. For the SRM portion of the self-efficacy questionnaire, Figure 46 depicts the posttest mean and standard error of the mean, while Figure 47 contains a box plot for both conditions.

Subhypothesis 3a: Table 15 contains the results of the independent samples t-tests performed on the averaged ratings for both the BRS and SRM items. Significance was found for the BRS averaged rating between groups, $t\ (34) = 2.64$, $p = .01$. Thus, condition was significantly related to BRS rating on the self-efficacy measure. An eta squared of .07 indicates that 7% of the variability in the BRS self-efficacy is related to differences in condition. The SBT group had higher mean ratings on both portions of the self-efficacy questionnaire. Therefore, participants in the SBT condition had significantly higher perceived levels of BRS self-efficacy than participants in the CBT condition, and subhypothesis 3a was supported.

Subhypothesis 3b: Significance was not found for the SRM self-efficacy questionnaire between groups, t (34) = 1.60, p = .06. Therefore, the SBT and the CBT groups had the same perceived levels of self-efficacy for the SRM portion of the SEQ and subhypothesis 3b was not supported.

***Table 14.* Means and Standard Deviations**

Table 14

Means and Standard Deviations for BRS and SRM Averaged Self-Efficacy Ratings

	n		*M*		*SD*		*St Error Mean*	
Performance Measure	SBT	CBT	SBT	CBT	SBT	CBT	SBT	CBT
BRS average	18	18	4.42	4.13	.26	.39	.06	.09
SRM average	18	18	4.18	3.87	.68	.49	.16	.11

***Table 15.* T-test**

Table 15

Independent Samples T-test Results for Comparison of SBT and CBT Means for BRS and SRM Self-Efficacy Measure

	95% CI					
	LB	UB	t	*df*	*p*	η^2
BRS average	.07	.52	2.64	34	.01	.07
SRM average	-.09	.71	1.60	34	.06	.05

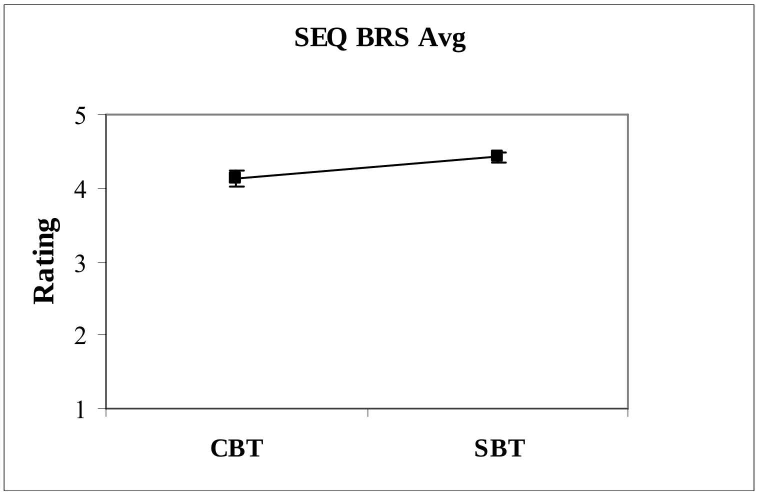


Figure 44. Mean and standard error of the mean for the BRS portion of the self-efficacy questionnaire, with rating as a function of condition.

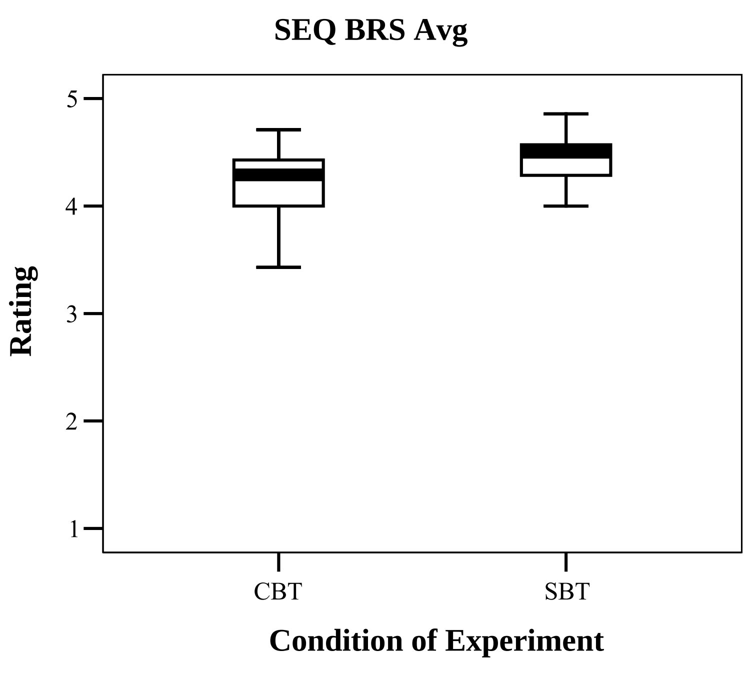


Figure 45. Box plot for the BRS portion of the self-efficacy questionnaire, with rating as a function of condition.

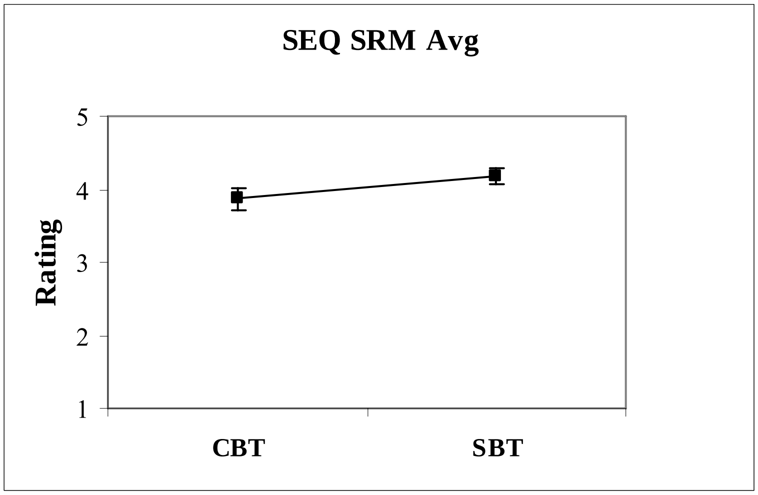


Figure 46. Mean and standard error of the mean for the SRM portion of the self-efficacy questionnaire, with rating as a function of condition.

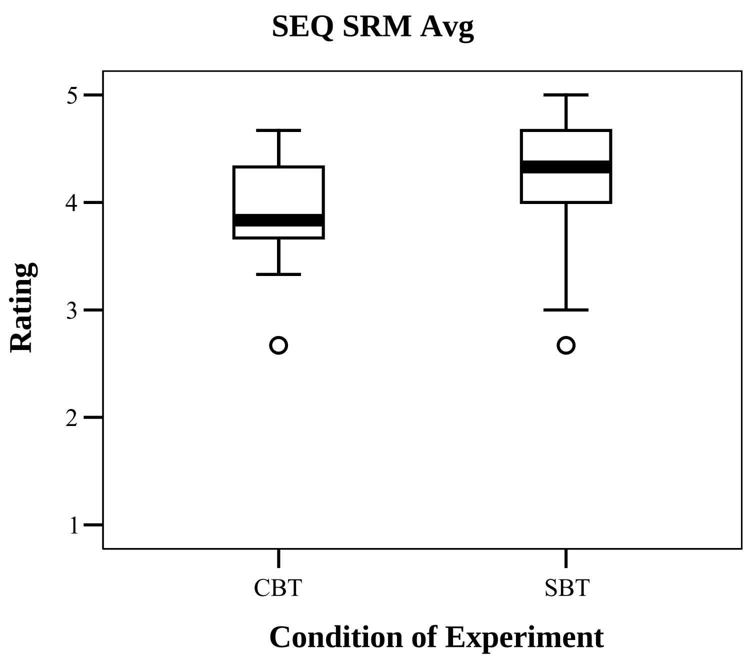


Figure 47. Box plot for the SRM portion of the self-efficacy questionnaire, with rating as a function of condition.

Subjective Workload

Hypothesis 4: Participants in the SBT condition will exhibit significantly lower levels of perceived workload than participants in the control condition. To test this hypothesis, differences in workload between groups and within test phases were analyzed using a two way ANOVA. Ratings for the six question NASA TLX subjective workload measure were averaged and compared between groups and within pre and post test sessions. Means and standard deviations for the TLX averaged ratings are reported in Table 16. Figure 48 depicts the mean and standard error of the mean for session and condition of the NASA TLX. Figures 49 and 50 contain box plots for the pre and posttest. For this measure of workload, a lower rating is actually better as it means the participants are less likely to become overwhelmed by the task and/or perform poorly.

Table 17 contains the results for the two-way ANOVA performed on the averaged ratings for the pre and posttest workload data. No significant differences were found for the pre and posttest comparison, $f(1, 34) = .81$, $p = .38$. Significant differences were also not found for the comparison between groups, $f(1, 34) = .02$, $p = .89$. Likewise, a significant interaction was not present, $f(1, 34) = .16$, $p = .69$. Therefore, the SBT group did not report a lower level of workload than the CBT group and hypothesis 4 was not supported.

***Table 16.* Means and Standard Deviations**

Table 16

Means and Standard Deviations for NASA TLX Averaged Ratings

Performance Measure	Test	*n*		*M*		*SD*	
		SBT	CBT	SBT	CBT	SBT	CBT
Averaged Rating	Pre	18	18	4.51	4.42	1.07	.75
	Post	18	18	4.32	4.35	.75	.75

***Table 17.* ANOVA**

Table 17

Two way Analysis of Variance for Pre and Post NASA TLX

Source	*df*	F	η^2	Power	*p*
Session (S)	1	.81	.02	.14	.38
Group (G)	1	.02	.00	.05	.89
S × G	1	.16	.00	.07	.70
S within-group error	34	(.35)			

Note. Values enclosed in parentheses represent mean square errors.

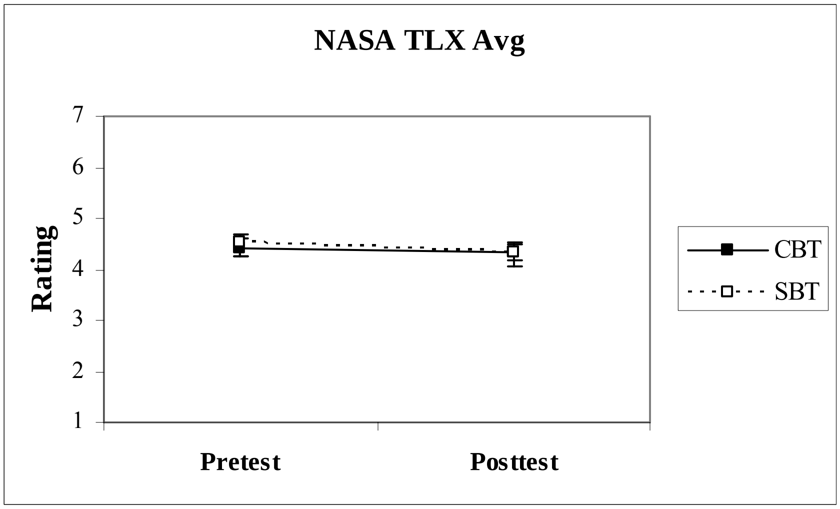


Figure 48. Mean and standard error of the mean for the NASA TLX, with rating as a function of session and condition.

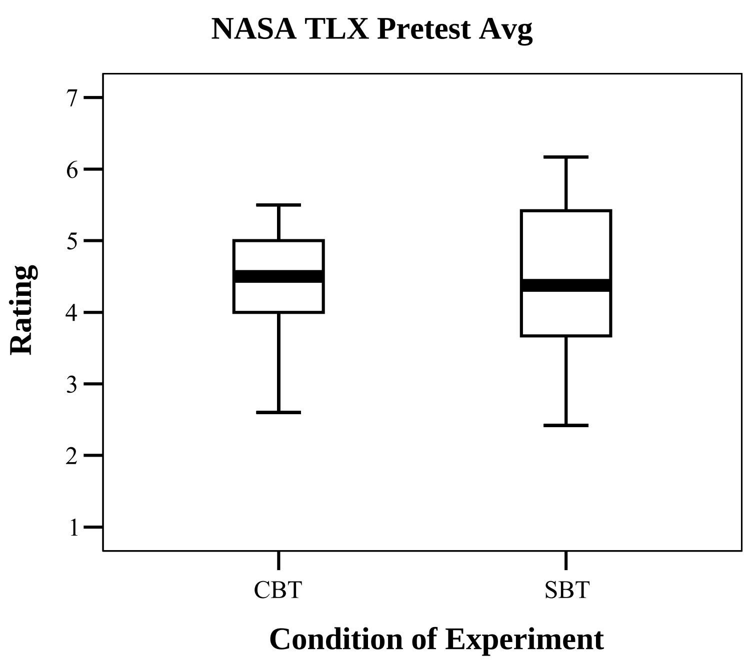


Figure 49. Box plot for the NASA TLX pretest, with rating as a function of condition.

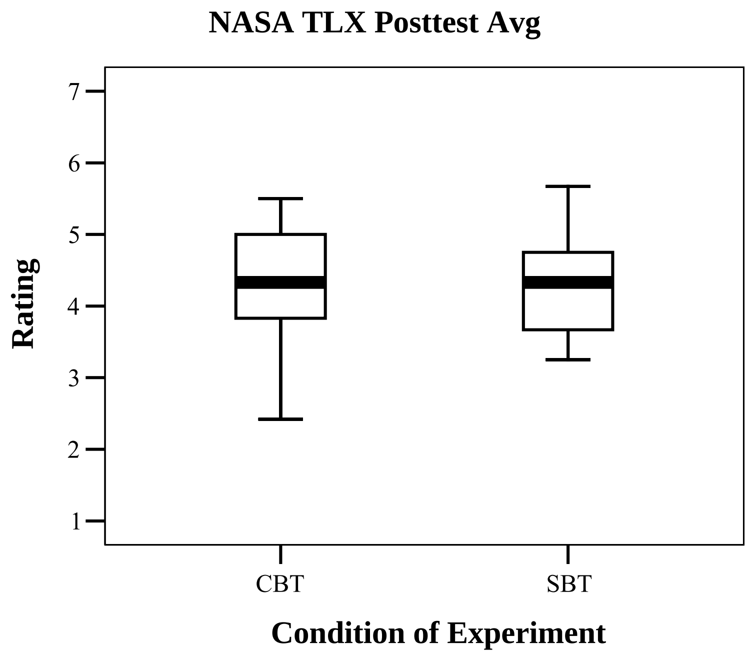


Figure 50. Box plot for the NASA TLX posttest, with rating as a function of condition.

DISCUSSION

The purpose of this study was to evaluate the effectiveness of scenario-based training as a method to train parachute use, and to compare SBT to traditional training. Certain KSA's were thought to be important in training effective use of the BRS parachute, such as situational awareness, decision-making, and knowledge of how to use the parachute. Scenario-based training utilizes practice and feedback with simulated scenarios to enhance performance, automate complex tasks, and develop knowledge and skills, such as decision making (Oser, Cannon-Bowers, Salas, & Dwyer, 1999; Oser, 1999). Indeed, SBT has been found successful, reducing response time, enhancing high order and procedural skills, reducing training costs and time, and has been rated as highly effective by former SBT trainees (Bowers & Morgan, 1991; Scharr, 2002; Lowry, 2000; Stewart et al., 2002). Therefore, SBT was anticipated to be a more effective training method for the BRS parachute than traditional training. The results for the effectiveness of training on performance are discussed below.

Support for Using Scenario-Based Methods to Train the BRS Parachute

In the current study, evidence supporting and not supporting SBT as a more effective method was found. The hypotheses and subhypotheses found to support SBT as a more effective method than traditional methods are discussed next.

Pilot Performance

Scenario-based training was found to be more effective than traditional training for ten of the 16 performance measures in this study. Related performance measures are grouped together in this part of the discussion, categorized by the underlying constructs they are believed to measure.

Performance measures related to BRS Deployment Decision

The performance measure “Pilot makes appropriate controlled landing/BRS decision” (subhypothesis 1a) was decomposed to provide a more detailed level of analysis and address the two major concerns with the BRS parachute: failing to look for a place to land before deploying the parachute and neglecting to use the parachute when necessary. Since the posttest involved scenarios and decisions that were closely related to scenarios and decisions practiced in the scenario-based training intervention, this practice was expected to enable participants in the SBT condition to make a good decision about parachute use more rapidly via pattern matching. This is in agreement with the RPD model of naturalistic decision making (Klein, 1989). Indeed, scenario-based training is believed to have the potential to improve decision making (Oser, et al. 1999). Thus, the finding of a significant effect was not surprising for the performance measure “Percent looked for a place to land before using the BRS parachute” (subhypothesis 1b). SBT participants tended to look for a place to land before they used the parachute more often than CBT participants. This measure reflects one of the main issues with the BRS decision, and is important mainly because of cost. Although saving lives is the primary concern with the BRS parachute, an important and related issue is the concern that pilots will use the parachute unnecessarily. Using the BRS without first attempting to land (when appropriate) could mean that the pilot will use the parachute unnecessarily. This is a problem because deploying the parachute results in high expenses for the pilot and/or insurance company, as well as to the government for investigating the accident.

The performance measure “Pilot uses parachute at correct time based on the sequence of events”, also referred to as “BRS timing” (subhypothesis 1h), was particularly important for parachute use, as timing is an essential aspect of effective decision making. In some scenarios, a correct decision (e.g. releasing the parachute) made too late might result in terrain collision (and often did). On the other hand, making the decision to use the parachute too early might mean that other alternatives were not utilized. If these other options *were* utilized, then the pilot might have been able to land safely without using the parachute. Releasing the parachute usually results in very costly damage to the plane, and more precise timing might mean the pilot is able to land in some emergencies without using the parachute and thus save thousands of dollars. As with the previous decision related performance measure, the practice received by SBT participants was expected to enable them to make a decision more rapidly via pattern matching. Therefore, the finding of a significant effect was likewise not surprising for this performance measure. Specifically, pilots in the SBT condition performed significantly better than pilots in the CBT condition. A look at the means for the two groups shows large differences in timing. SBT participants generally used the BRS parachute within an acceptable amount of time, while the CBT participants often used the parachute too early or too late. Performance with the parachute was extremely poor in the pretest; few participants used the parachute. Thus, no pretest rating could be assigned for this and other BRS related measures.

The performance measure “Pilot uses parachute above the minimum altitude (500 ft)” or “BRS altitude” (subhypothesis 1i) was strongly related to the measure of timing discussed in the previous paragraph. Both measures assessed *when* the BRS parachute was deployed. However, in the "above minimum altitude" measure, the focus was whether or not the pilot was above the minimum altitude when the decision was made. In some cases a pilot might make the decision too late or early and yet still be above the minimum altitude. Therefore, two separate performance measures for the similar concept were warranted. Since the "timing" and the "above minimum altitude" measures were assessing similar concepts, significance for both measures was not surprising. Participants in the SBT condition again performed significantly better than participants in the CBT condition on the "above minimum altitude" performance measure. Both groups of pilots generally used the parachute above 500 ft, but the CBT group failed to use the parachute above 500 ft more frequently than did the SBT group. Explanations of significance discussed in the BRS timing section are applicable here as well. Pretest data was not available for this performance measure as very few participants used the parachute in the pretest.

Performance Measure Related to SRM

Significance was anticipated for the performance measure "Pilot utilizes 5 P/SRM technique appropriately" (subhypothesis 1e), as the SBT group received more SRM training than the CBT group. SRM was incorporated within the training intervention for the SBT condition with the idea that using this technique might help pilots even more with SRM related behaviors such as situational awareness, decision making, and risk management (FAA/Industry Training Standards, 2004). If pilots did not learn or utilize the technique as effectively as hoped, this might have resulted in less than anticipated significance for the performance measures related to decision making, risk management and situational awareness. Conversely, use of the 5 P technique might enhance skills like decision making for SBT participants (even if the technique is not used perfectly). Indeed, it may have been responsible for part of the significant effect noted in the BRS timing discussion, as timing is considered an important part of the BRS deployment decision.

The interaction was significant for the SRM measure. Specifically, SBT pilots improved more from pre to post test than did the CBT pilots. The means for the CBT group and SBT group in the pretest were both equal to one, indicating that pilots were not using SRM (1 = "very inappropriate use" on the rating scale). While the mean for the SBT posttest was significantly higher, they still indicate the participants were not using SRM as often as appropriate.

There are a couple of possible explanations as to why the 5 P's technique was not utilized more frequently by SBT participants. The training time might have been too brief for some participants to feel comfortable using the 5 P's technique. Similarly, the training session might have covered too much other information. Perhaps a training session devoted only to SRM would result in more frequent use of the SRM technique by pilots. It is important to note that although the SBT participants did not perfect their SRM skills in this study, they did show significant improvement with very little training time devoted to SRM.

General Emergency Performance Measure Related to Skills

The scenario-based training methodology trains within a realistic and well-integrated environment (Oser, et al. 1999). Because of this integrated practice environment, a variety of emergency related behaviors and skills (e.g. maintaining control of aircraft despite strong winds or other skills) were practiced to some degree.

During the practice phase of the SBT training, participants received feedback not only on their performance regarding BRS related performance measures, but also on their performance with general emergency procedures. Thus, since these behaviors were also addressed in the SBT scenarios, it was anticipated that the SBT pilots would perform better than the CBT pilots on these behaviors.

Both groups performed well on the measure “Pilot maintains control of the aircraft” (subhypothesis 1k). Mean ratings were close to five, the highest rating possible on the scale ranging from 1-5. Despite the overall effective performance in both groups, participants in the scenario-based training condition still performed significantly better than participants in the computer based training condition. The significant difference between groups may indicate that pilots in the SBT group were better prepared for the emergencies in general and were not as flustered or distracted by problems (e.g. severe winds) encountered in flight. Therefore, maintaining control of the aircraft was not as difficult for them. This would support the idea that SBT improves performance (Oser, et al. 1999). Another possibility is that pilots in the SBT condition simply knew what to expect during the posttest better than CBT pilots and therefore could react more effectively.

General Emergency Performance Measures Related to Procedural Knowledge

Participants in the SBT condition were expected to develop relevant skills and procedural knowledge from the practice and feedback within the simulated scenarios (as described by Oser, et al. 1999). Thus, SBT participants were anticipated to recall emergency checklist procedure more effectively than CBT participants for the performance measure “Pilot follows checklist procedure to resolve problem” (subhypothesis 1m). Participants in the SBT condition performed significantly better than participants in the CBT condition on this measure. The SBT group had a mean rating for this measure which indicated that many followed the checklist procedure (at least for some of the steps). The CBT group had a mean rating that indicated that pilots failed frequently to follow the checklist procedure.

Most likely, the improved performance of the SBT group can be attributed to the SBT approach. However, it is unclear whether participants in the CBT condition would actually fail to follow the checklist in actual flight. In other words, it may be that these participants were simply not treating the simulation like an actual flight. Additionally, the low mean ratings for both groups may indicate that pilots are not always following emergency procedure in actual flight.

Another performance measure relating to procedural knowledge was “Pilot contacts ATC” (subhypothesis 1n). Based on the SBT rationale (Oser et al., 1999), participants were expected to contact ATC when appropriate more often than CBT participants. Participants in the SBT condition did perform significantly better than participants in the CBT condition on this measure. The SBT group had a mean rating for this measure that indicated generally effective contact with ATC, while the CBT group had a mean rating that indicated generally ineffective contact with ATC. Timing was taken into consideration with this measure. Participants who contacted ATC at the right time (generally the earlier the better) received higher ratings than participants who did not. The precise effect of timing on performance is unclear for this measure. Overall, the mean ratings for SBT participants indicated that they generally contacted ATC within an acceptable amount of time. On the other hand, the mean rating for CBT participants indicates they either did not contact ATC, or if they did, it was generally later than necessary.

Since pilots in the SBT condition had more practice in handling related emergencies and performed better as a result, it is likely that this difference probably was caused by the SBT method. As was previously discussed, however, it is possible that some of the CBT participants were not treating the simulation device like an actual flight. Pilots may have communicated with the pilot acting as ATC in this experiment differently than they would communicate with ATC in real life. The additional practice received by participants in the SBT condition might have made them feel more comfortable communicating with ATC in a controlled setting, and thus increased the frequency they contacted ATC during the post-test. Assuming the SBT intervention was responsible for the performance improvement, the implication is that refresher courses in emergency procedures could be beneficial to pilots throughout their career, particularly if scenario-based training is used.

Additionally, the performance measure “Pilot declares an emergency” (subhypothesis 1o) was another procedural measure. Participants in the SBT condition performed significantly better than participants in the CBT condition for this measure. That is, SBT participants generally declared an emergency within an acceptable amount of time. On the other hand, the mean rating for CBT participants indicated that they either did not declare an emergency, or if they did, it was generally later than necessary. Declaring emergencies and contacting ATC are often closely related, so it is appropriate both were found significant. Thus, the previous discussion of contacting ATC may be referred to for this measure.

The performance measure “Pilot diverts or continues to destination around storm” (subhypothesis 1p) also involves procedural knowledge, and, again, the SBT group was expected to perform more effectively than CBT for this performance measure (i.e., Oser, et al. 1999). Participants in the SBT condition did perform significantly better than participants in the CBT condition on this measure. The SBT group had a mean rating for this measure that indicated the pilots usually diverted when appropriate and at an acceptable point in time, while the CBT group had a mean rating that indicates they were less inclined to divert when necessary. Again, participants in the CBT condition might perform differently in actual flight if they were not treating the simulation device like a real aircraft. However, since pilots in the SBT condition had additional practice in handling related emergencies and performed more effectively as a result, this significant difference probably indicates that refresher courses in emergency procedure might help to lower the rate of aviation accidents.

Overall Emergency Performance Measure

Scenario-based training was anticipated to better prepare participants in the SBT condition overall, as it allows participants to practice and enhance many important behaviors. The situated practice and feedback received via simulated scenarios are expected to enhance overall performance, including higher order skills and procedural knowledge (Oser, et al. 1999). The performance measure “Overall the pilot responded” (subhypothesis 1f) was intended to capture this overall performance. Raters were instructed to take all behaviors during each emergency event into consideration when assessing overall performance. Both groups improved in the posttest, but participants in the SBT had higher ratings of overall performance than did participants in the CBT group, as expected.

Knowledge Test

SBT is believed to effectively develop procedural knowledge via practice in simulated scenarios (Oser, et al. 1999), so significance was expected for both portions of the knowledge test. This was true for the SRM portion of the knowledge test (subhypothesis 2b). Significance was not surprising, as the SBT group received more instruction on SRM and was encouraged to use it more frequently. Although the SBT group performed significantly better, they were often unable to answer one of the questions on the knowledge test (i.e. list the six behaviors of SRM). However, if additional training time had been available for SRM, it is highly likely that the SBT participants would have scored higher on average.

Self-Efficacy

Since a high level of self-efficacy is positively correlated with performance (Bandura, 1997), it was anticipated that the SBT methodology would foster higher degrees of self-efficacy ratings than would the CBT condition. Participants in the SBT condition did have significantly higher self-efficacy than participants in the CBT condition on the BRS related questions (subhypothesis 3a). As noted earlier, the SBT group performed better than the CBT group on many of the performance measures as well. Therefore, it is appropriate that they also had higher self-efficacy for the BRS parachute.

Lack of Support for the Effectiveness of SBT

Although SBT participants performed better than CBT participants on several of the measures, not all measures showed SBT as a more effective method. These measures are discussed next.

Pilot Performance

Scenario-based training was not found to be significantly different from traditional training for six of the 16 performance measures in this study. Related performance measures are grouped together in this part of the discussion, categorized by the underlying constructs they are believed to measure (i.e. decision making, general emergency procedures, and BRS procedures).

Performance Measures Related to the BRS Deployment Decision

Scenario-based training is believed to be an optimal method for training decision making because it offers more practice in realistic situations (Oser, et al. 1999). SBT in this experiment offered practice via emergency events involving decisions similar to those the pilot would make in the posttest. This should have enabled SBT participants to make more rapid decisions via pattern-matching, similar to the way experts make decisions according to the Recognition-Primed Decision model (Klein, 1989). Also, the 5 P technique utilized by SBT participants is believed to enhance SRM behaviors such as decision making (FAA/Industry Training Standards, 2004). The SBT group was therefore expected to improve more from the pretest to the posttest for a variety of performance measures. This improvement, however, did not occur for the performance measure “Pilot makes appropriate controlled landing/BRS decision” (subhypothesis 1a).

There are several possible explanations for this lack of improvement. One is that the number of participants could have been a little too small; a few more participants might have easily made it significant. This is supported by the finding that power had only a .49 chance of detecting a significant difference at the .05 level.

Alternatively, the problem might lie in the training, the participants, or with the performance measure itself. Perhaps the most plausible reason is the performance measure itself, as the measure was at times ambiguous to the raters. While the wording of this measure was applicable in some of the scenarios, it was confusing in other scenarios. This measure was designed to assess the effectiveness of the pilot's decision to deploy the parachute or land the plane. However, confusion about the measure early in the experiment might have resulted in some inappropriate ratings.

Despite the lack of significance for the interaction, the majority of participants made a poor decision in the pretest, choosing to land in the pretest when they should have used the parachute. Although participants were told what the parachute was and where it was located in the simulation device before the pretest flight, many admitted to completely forgetting about the parachute when the emergency event occurred in the pretest. This indicates a need for training; simple exposure to the parachute option is not enough to use the parachute in an emergency.

As noted previously, the performance measure "Pilot makes appropriate controlled landing/BRS decision" was subdivided into two parts. The performance measure "BRS when necessary" (subhypothesis 1c) reflects one of the main issues with the BRS parachute, namely, that pilots will not use the parachute when absolutely necessary to prevent a potentially fatal aviation accident. This is a growing concern in the general aviation industry, as a high number of accidents recently have caused the FAA to reassess current training standards (FAA/Industry Training Standards, 2004). As with the other performance measures that involve decision making, the practice provided by SBT was expected to improve participant's decision making abilities for this performance measure (Klein 1989; Oser, et al. 1999). This improvement, however, did not occur for the performance measure "BRS when necessary".

The lack of significance probably indicates that pilots in the both groups will perform similarly in a real plane regarding using the parachute when necessary. Alternatively, there may have been an issue with the training or with the performance measure itself.

In addition, the performance measure "Frequency did not crash" (subhypothesis 1g) was included to further assess decision making. Participants in the SBT condition were expected to crash less frequently than participants in the CBT condition. This was anticipated mainly because SBT is believed to be an effective method for decision making (Oser, et al. 1999), and making a poor decision regarding landing or using the BRS should sometimes result in crashing. Overall, participants in the SBT condition did not perform significantly better than participants in the CBT condition for this measure. There are a few possibilities for this lack of significance. The lack of significance for this measure may have been because CBT pilots were slower to make the decision to use the BRS or land, but still made the correct decision in time to avoid crashing. Or perhaps pilots in the CBT group were more often deploying the parachute too early. In this case, crashing would not be a problem, but early deployment of the parachute could instead result in unnecessary damage to the aircraft. Another possibility lies in the level of analysis. That is, the results might have been different if the groups had been analyzed by event. A combined rating that took all events into account was used for the "Did not crash" percentage, yet in some events crashing was highly unlikely and therefore it might have been more appropriate to analyze each individually.

Performance Measures Related to BRS Procedures

Oser et al. (1999) emphasizes the practice and corrective feedback of SBT as being particularly useful for the development of skills and procedural knowledge (such as the steps for using the parachute). Thus significance for the performance measure "Pilot uses parachute correctly", also referred to as "Correct BRS use" (subhypothesis 1d), was anticipated, as this measure most strongly reflects the procedural knowledge required to deploy the BRS parachute. However, a significant effect was not found for the interaction of condition and session for the performance measure "Correct BRS use". As the majority of participants were unfamiliar with the BRS parachute, participants in both conditions rarely used the parachute in the pretest, and no one used it correctly. In the posttest both groups frequently used use the parachute, but at times used it incorrectly. The lack of difference between groups on the post-test was really not surprising for this measure, as both groups received a fair amount of training and practice specifically on the deployment procedure. The CBT group may have even received more training on this skill than the SBT group, as they used a computer based simulation which involved watching a demonstration of the procedure and then repeatedly practiced the procedure on their own. This computer based simulation was unavailable to participants of the SBT condition. Thus, all participants received a fair amount of practice, which likely helped to automate the procedure for both groups in the posttest.

Significance was anticipated for the performance measure “Pilot uses parachute below the maximum rate/knots” or “BRS knots” (subhypothesis 1j), as participants in the SBT condition received practice and feedback within simulated scenarios, and thus would develop procedural knowledge (Oser, et al. 1999). This should have resulted in SBT participants recalling to use the parachute below the maximum rate more often than CBT participants. However, a significant difference between groups was not found for this measure. Even so, the majority of participants in both groups received high ratings for this performance measure, indicating both groups were below the maximum rate most of the time. Therefore, remembering to deploy the parachute below the maximum rate did not appear to be difficult. Pretest data was not available for this performance measure as pilots generally did not use the parachute in the pretest.

General Emergency Performance Measure Related to Procedural Knowledge

SBT is believed to be a successful method for enhancing performance with a tight linking of objectives, scenarios, performance measures and feedback (Oser, et al. 1999). Thus, participants in the SBT condition were expected to perform significantly better on all measures, including the measure “Pilot refers to checklist” (subhypothesis 1l). However, participants in the scenario-based training condition did not perform significantly better on this measure than participants in the computer based training condition. Both groups performed ineffectively. Participants in both conditions performed more poorly on checklist related performance measures than they did on all other measures for general emergency procedures.

These findings indicate that perhaps pilots forgot the checklists were available or they were reluctant to use them for some reason. It is unclear whether participants would take the same inaction with the checklist in actual flight, or if they were simply not treating the simulation device like a real plane as instructed. The surprisingly low means of participants could also indicate a problem with the measure (e.g. some situations did not really require looking at the checklist – pilots have many sections of the checklist memorized). Raters were told to take this into account, but it is still possible the ratings were adversely skewed for this measure. The checklist performance measures were also used to assess whether pilots noticed a few of the more minor emergency events, such as alternator problems. Some pilots who might otherwise have used the checklist might not have even noticed the problem. Therefore, using the checklist would not apply.

Knowledge Test

Although SBT is believed to effectively develop procedural knowledge through practice (Oser, et al. 1999) and significance was found for the SRM knowledge test, both conditions performed about the same on the BRS knowledge test (subhypothesis 2a). Although practice might have been helpful for SBT participants during the performance assessment, it did not appear to make a difference with BRS knowledge. Furthermore, both conditions received the same information about the parachute; the main difference between groups was only the way the information was received (i.e. reading vs. practice). Written knowledge tests might be more appropriately matched with other forms of instruction, such as reading or lectures. What is interesting is that the CBT group had an equivalent amount of knowledge (as assessed by this knowledge test), yet still did not perform as effectively as the SBT group on many BRS performance measures.

Self-efficacy

As discussed previously, a high level of self-efficacy is positively correlated with performance (Bandura, 1997). Therefore, it is unclear why the SBT condition did not rate higher on the SRM related measures (subhypothesis 3b). The SBT condition had both higher ratings on the SRM performance measure and higher scores on the knowledge test for SRM. One possible explanation might be that although the SBT group performed significantly better than the CBT group, neither group performed well (e.g. both groups received a mean rating that reflected, at best, an ineffective use of the SRM technique on the posttest). Many SBT participants might not have been fully comfortable with the technique, and thus did not have significantly higher self-efficacy regarding SRM. The self-efficacy questionnaire was administered after the posttest (as opposed to after training), so it is difficult to say if efficacy affected performance or if performance affected efficacy.

Workload

Significance was expected for both condition and session, as workload is one stressor (Wickens, Gordon, & Liu, 1998) and a high amount of workload can negatively impact performance. However, significance was not found in the interaction of session and condition for the averaged ratings of the NASA TLX. Participants in SBT were expected to perceive less workload due to the extensive practice provided by SBT. This, in turn, could imply they have more mental resources available to make effective decisions regarding the parachute. The lack of significance most likely suggests that SBT does not reduce perceived workload. Performance and workload are most likely not related for use of the BRS parachute, at least not according to the results of this study.

Additionally, the lack of results might indicate that the measure was not sensitive enough or administered properly. Using a different measure of workload might have resulted in different findings.

In summary, SBT is believed to effectively develop skills and procedure knowledge. Thus, participants in the scenario based training condition were expected to perform significantly better than participants in the traditional training condition. This was true for ten of the 16 performance measures. However, six of the performance measures, including four BRS related measures, were not significant. Overall, most of the hypotheses for this study were at least partially supported (see Table 18 for a list of supported/unsupported hypothesis). Despite the lack of significance for some measures, SBT participants never performed or scored significantly lower than CBT participants.

Additionally, when assessing the effectiveness of the training intervention it is important to note that the study took place over a two day period, with a day between the training and testing sessions. Therefore, the significant improvements found for session and condition provides some indication that the training effects are retained over time. Overall, SBT appears to be an effective method for enhancing performance with the BRS parachute and is probably more effective than CBT based on the results from this study.

Table 18. **Summary of Support for Hypotheses/Subhypotheses**

Hypotheses/Subhypotheses that were supported (significant at the .05 level)
1b: Participants in the SBT condition will perform more effectively than will the participants in the control condition for the subhypothesis "Looked for place to land"
1e: "SRM"
1f: "Overall performance"
1h: "BRS timing"
1i: "BRS altitude"
1k: "Maintains control of aircraft"
1m: "Follows checklist"
1n: "Contacts ATC"
1o: "Declares emergency"
1p: "Diverts"
2b: Participants in the SBT condition will achieve significantly higher scores on the SRM portion of the knowledge test than will participants in the control condition.
3a: Participants in the SBT condition will have significantly higher ratings of self-efficacy than the control condition on the BRS portion of the Self-Efficacy Questionnaire.
Hypotheses/Subhypotheses that were not supported (at the .05 level)
1a: Participants in the SBT condition will perform more effectively than will the participants in the control condition for the subhypothesis "Controlled landing/BRS decision"
1c "BRS when necessary"
1d: "Correct BRS use"
1g: "Frequency did not crash"

1j: "BRS knots"
1l: "Refers to checklist"
2a: Participants in the SBT condition will achieve significantly higher scores on the BRS portion of the knowledge test than will participants in the control condition.
3b: Participants in the SBT condition will have significantly higher ratings of self-efficacy than the control condition on the SRM portion of the Self-Efficacy Questionnaire.
4: Participants in the SBT condition will exhibit significantly lower levels of perceived workload than participants in the control condition.

CONCLUSION

The BRS parachute is being used increasingly in aircraft and requires pilots to make an additional decision under extreme pressure. This study examined the effect of training on the decision of deployment. The training goal was to equip pilots with the knowledge and skills to deploy the chute when necessary and avoid unnecessary deployment. Although airplanes are equipped with the parachute to enhance survivability, some researchers and insurance companies believe that this option may cause the unnecessary destruction of aircraft. Pilots must know how and when to use the parachute to avoid unnecessary losses of lives or of expensive equipment. Using SBT to train parachute use offered trainees realistic, challenging scenarios to practice using the skills and knowledge essential for effective parachute use. Pilots trained with SBT tended to perform better than those trained with traditional methods, as shown in Table 18. Specifically with decision making, the results showed that pilots trained with SBT more often made the decision to use the BRS at the appropriate time, and did not use the parachute without at least looking for a place to land first. SBT pilots also tended to have higher self-efficacy for BRS use, and scored higher on a knowledge test for SRM.

Future Recommendations

Participants trained with SBT might have performed even better if they had longer training time. Due to budget and time constraints, the training time was limited to three hours. The large amount of material packed into the scenario-based training intervention might have caused some pilots to feel overwhelmed and to lose focus. Some pilots might have forgotten some of the material over the course of the two day training and testing period as a result of the potentially overwhelming amount of new information. Additional training time might help pilots recall new material when being tested within the experiment as well as later on in actual flight. Additional time could also mean the inclusion of more training scenarios, which might help pilots perform better in a greater variety of emergency situations.

Another issue with this research was the wide range of flight hours participants had, from 100 to 300. While all pilots seemed to benefit from training, it is difficult to compare a pilot with 100 flight hours to a pilot with close to 300 hours. Pilots might have been divided into two groups (i.e. 100-200, 201-300) or the range might have been narrowed to only include one of those possible groups. Also, some anecdotal evidence was noted regarding a difference in performance for participants trained at ERAU compared to those who received the majority of their flight training at other locations. A future study might look at these two groups as another independent variable or just look at ERAU pilots.

Future research might also involve the general emergency aspect of this study. It appears that pilots might need a refreshing training course in handling emergencies effectively, as some pilots failed to follow proper emergency procedure. Many pilots, particularly those in the CBT group, did not use checklists, failed to contact ATC and/or did not declare an emergency. Pilots in training, even at an excellent flight training program such as Embry-Riddle's, might benefit from recurrent training courses in emergency management.

To see if the effects of the scenario-based training intervention will last over time, another similar study might be performed with additional testing sessions spread out over several months or even years. Otherwise it is difficult to say whether the training is truly beneficial in the long term. A longitudinal study could also include analyses of aviation accident statistics to see if scenario-based training is beneficial in real life. Namely, do pilots who have received SBT handle a real emergency more effectively than other pilots? Is scenario-based training in emergencies linked to a reduced accident rate? However, many pilots would have to receive scenario-based training in emergency situations in order for this study to be possible. The effectiveness of refresher courses in emergency procedure could also be analyzed in this way.

SRM was one of the goals within this study and was taught mainly to SBT participants. Training SRM was somewhat successful, as pilots who received additional training on SRM and the 5 P's performed significantly better on the posttest SRM performance measures and SRM knowledge test than those who did not receive as much SRM training. However, assessment of the means of the SRM measures indicates pilots may need more extensive SRM training than what was provided in this study. This is further supported by the lack of a significant difference in SRM self-efficacy between groups. A training intervention devoted solely to SRM might improve performance on such related measures, as would additional training time.

As noted previously in the discussion, the significantly higher ratings for the SBT condition on many of the measures likely indicates that SBT was an effective method and refresher courses in emergency procedure might be beneficial to pilots throughout their career. However, the higher rating of SBT participants could indicate that although pilots do need more extensive training, it is only on specific scenarios. For example, the icing scenario in this experiment confused many participants, as most have never been exposed to icing (neither in real life nor in a simulation device). Pilots would probably handle an unfamiliar emergency such as this more poorly than one they have experienced before. This is supported by the Recognition-Primed Decision theory of pattern matching by expert decision makers (Klein, 1989). Thus having received training specifically for a scenario with icing might have resulted in better scores for all participants. Regardless, it appears the scenario-based training was beneficial to participants.

In summary, both types of training were an improvement over no training. However, the scenario-based training appeared to be overall more effective than the traditional training on a variety of measures. Additional research is recommended so more conclusive results may be obtained regarding the effectiveness of scenario-based training on single-pilot resource management and the Ballistic Recovery System.

REFERENCES

Aamodt, M.G. (2004). *Applied Industrial/Organizational Psychology*. Belmont, CA: Wadsworth/Thompson Learning.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: W.H. Freeman

Bandura, A. (2000). Cultivate self-efficacy for personal and organizational effectiveness. In E. A. Locke (Ed.), *Handbook of principles of organization behavior* (pp. 120-136). Oxford, UK: Blackwell.

Bell, B., & Mauro, R. (2000). Training in judgment and aeronautical decision-making. In *Proceedings of the 19th Digital Avionics System Conference* (pp. 5.B.1-1-5.B.1-8). Piscataway, NJ: Institute of Electrical and Electronic Engineers.

Benner, L. (1975). D.E.C.I.D.E. in the hazardous materials emergencies. *Fire Journal, 69* (4), 13-18.

Bonsor, K. (n.d.) *How ejection seats work.* Retrieved March 2, 2005, from http://www.howstuffworks.com/ejection-seat1.htm

Bowers, C.A., & Morgan, B.B. (1991). *Simulator scenario development for aircrew coordination training.* Technical report submitted to Naval Training Systems Center, Orlando, FL.

Callaghan, K.S., & Irwin, R.J. (2001). The decision to eject: A receiver operating characteristic analysis. *Aviation, Space, and Environmental Medicine, 72*(11), 1017-1024.

Callaghan, K.S., & Irwin, R.J. (2003). Ejection performance of strike pilots: Effect of the designated decision height. *Aviation, Space, and Environmental Medicine 74*(8), 833-837.

Cannon-Bowers, J.A., Burns, J.J., Salas, E., & Pruitt, J.S. (1998). Advanced technology in scenario-based training. In J.A. Cannon-Bowers & E. Salas (Eds.), *Making decisions under stress: Implications for individual and team training* (pp. 365-374). Washington, D.C.: APA.

Chase, W.G., & Simon, H.A. (1973). The mind's eye in chess. In W.G. Chase (Ed.), *Visual information processing* (pp.215-281). New York: Academic Press.

Cook, R.I., Woods, D., & McDonald, J. (1991). *Human performance in aesthesia: A corpus of cases* (CSEL91.003). Columbus, OH: Ohio State University, Cognitive Systems Engineering Laboratory.

Cook, R. I., & Woods, D. D. (1994). Operating at the sharp end: The complexity of human error. In M. S. Bogner (Ed.), *Human error in medicine* (pp. 255-301). Hillsdale, NJ: Earlbaum.

Driskill, W.E., Weissmuller, J.J., Quebe, J., Hand, D., Dittmar, M., & Hunter, D.R. (1997). The use of weather information in aeronautical decision-making (DOT/FAA.AM-97/3). Washington, DC; Federal Aviation Administration, Office of Aviation Medicine.

Driskill, W.E., Weismuller, J.J., Quebe, J.C., & Hand, D.K. (1998). Evaluating the decision-making skills of general aviation pilots (DOT/FAA/AM-98/7). Washington, DC: Federal Aviation Administration, Office of Aviation Medicine.

Dwyer, J.D., Oser, R.L., Salas, E., & Fowlkes J.E. (1999). Performance measurement in distributed environments: Initial results and implications for training. *Military Psychology, 11*(2), 189-215.

Endsley, M.R. (1988). Design and evaluation of situation awareness enhancement. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (Vol. 1, pp. 97-101). Santa Monica, CA: HFES.

Endsley, M.R. (2000). Theoretical underpinnings of situation awareness: A critical review. In M.R. Endsley, & D.J. Garland (Eds.), *Situation awareness analysis and measurement* (pp. 3-32). Mahwah, NJ: Erlbaum.

FAA/Industry Training Standards. (2004). *FITS curriculum recognition criteria.* Retrieved January 28, 2005, from http://learn.aero.und.edu/pages.asp?PageID=13814

Flightlogic EFIS synthetic vision flight display pilots operating guide and reference. (2004). Boise, ID: Chelton Flight Systems.

Fowlkes, J., Dwyer, D.J., Oser, R.L., & Salas, E. (1998). Event-based approach to training (EBAT). *The International Journal of Aviation Psychology, 8*(3), 209-221.

Gist, M. E. (1989). The influence of training method on self-efficacy and idea generation among managers. *Personnel Psychology, 42*, 787-805.

Goh, J., & Weigmann, D.A. (2002). Relating flight experience and pilots' perception of decision-making skill. In Proceedings of the Human Factors and Ergonomics Society 46th annual meeting. Baltimore, MD: HFES.

Goodman, C. (1998). Factors affecting the decision to eject. In *Proceedings of the SAFE 36th annual symposium* (pp. 492-500). Phoenix, AZ: SAFE association.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock, & N. Meshkati (Eds.), *Human Mental Workload* (pp. 139-183). Amsterdam: North-Holland.

Hays, R.T., Jacobs, J.W., Prince, C., & Salas, E. (1992). Requirements for future research in flight simulation training: Guidance based on a meta-analytic review. *The International Journal of Aviation Psychology, 2*(2), 143-158.

Hedge, J.W., Borman, W.C., & Hanson, M.A. (1996). Videotaped crew resource management scenarios for selection and training applications. In *Proceedings of the 38th Annual Conference of the International Military Testing Association*(TE36b). Retrieved on April 2, 2005 from http://www.ijoa.org/imta96/PAPER60.html

Hunter, D. (1997). Pilot characteristics. In R.A. Telfer & P.J. Moore (Eds.), *Aviation training: Learners, instruction and organization* (pp. 41-53). Brookfield, VT: Ashgate.

Jensen, R.S. (1982). Pilot judgment: Training and evaluation. *Human Factors, 24*, 61-73.

Karp, M.R. (2001). University aviation education: An integrated model. In *Proceedings of the 11th International Symposium on Aviation Psychology* (6 pp.). Columbus, OH: Ohio State University.

Karp, M., & Nullmeyer, R.T. (2001). Cockpit resource management for single-seat fighter pilots. In *Proceedings of the 11th International Symposium on Aviation Psychology* (7 pp.). Columbus, OH: Ohio State University.

Klein, G. (1989). Recognition-primed decisions. *Advances in Man-Machine Systems Research, 5*, 47-92.

Klein, G. (1997). Current status of NDM. In R. Flin, E. Salas, M. Strub, L. Martin (Eds.), *Decision making under stress*, (pp. 11-28). Hants, England: Ashgate.

Klein, G. (2000). How can we train pilots to make better decisions? In H.F. O'Neil, Jr. & D.H. Andrews (Eds.), *Aircrew Training and Assessment* (pp. 165-195). Mahwah, NJ: Erlbaum.

Klein, G.A., Orasanu, J., Calderwood, R., & Zsambok, C.E. (1993). *Decision making in action: Models and methods*. Norwood, NJ: Ablex.

Klein, G.A. (1998). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.

Kneebone, R.L., Nestel, D., Moorthy, K., Taylor, P., Bann, S., Munz, Y., et al. (2003). Learning the skills of flexible sigmoidscopy – the wider perspective. *Medical Education, 37*(1), 50-58.

Kocks, K. (2005). Systems that permit everyone to fly. *Aviation Today*. Found online March 15, 2005 from http://www.aviationtoday.com/cgi/av/show_mag.cgi?pub=av&mon=0301&file=0301sats.htm

Kraiger, K., Ford, J. K., and Salas, E. (1993). Integration of cognitive, behavioral, and affective theories of learning into new methods of training evaluation. *Journal of Applied Psychology Monograph, 78*, 311-328.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting and task performance*. Englewood Cliffs, NJ: Prentice- Hall.

Lauber, J.K. (1984). Resource management in the cockpit. *Air Line Pilot, 53*, 20-23.

Lauber, J.K., & Foushee, H.C. (1981). Guidelines for line-oriented flight training, volume 2 (NASA-CP-2814-VOL-2). Moffett Field, CA: National Aeronautics and Space Administration.

Loftin, R.B., Wang, L., & Baffes, P. (1989). Intelligent scenario generation for simulation-based training. In *AIAA Computers in Aerospace Conference, 7th, Monterey, CA* (pp. 581-588). Washington, DC: American Institute of Aeronautics and Astronautics.

Lowry, K.D. (2000). United States probation/pretrial officers' concerns about victimization and officer safety training. *Federal Probation, 64*(1), 51-55.

Medin, D.L., & Ross, B.H. (1992). *Cognitive psychology*. Orlando, FL: Harcourt Brace Jovanovich.

Mulder, J.A., van Paassen, M.M., & Mulder, M. (2004). Perspective guidance displays show the way ahead. *Journal of Aerospace Computing, Information, and Communication, 1*(11), 428-431.

NASA Facts Online. (2000). *Glass Cockpit Fact Sheet* (FS-2000-06-43-LaRC). Retrieved January 31, 2005, from http://oea.larc.nasa.gov/PAIS/Glasscockpit.html

Noe, R.A. (1999). *Employee Training & Development.* USA: The McGraw-Hill Companies.

O'Hare, D. (1990). Pilots' perception of risks and hazards in general aviation. *Aviation, Space and Environmental Medicine, 61*, 599-603.

O'Hare, D. (1992). The "artful" decision maker: A framework model for aeronautical decision making. *The International Journal of Aviation Psychology, 2*(3), 175-191.

Orasanu, J. (1997). Shared problem models and flight crew performance. In N. Johnston, N. McDonald, and R. Fuller (Eds.), *Aviation Psychology in Practice* (pp. 255-285). Hants, England: Ashgate.

Orasanu, J. & Connolly, T. (1993). In G. A. Klein, J. Orasanu, R. Calderwood, & C.E. Zsambok (Eds.), *Decision making in action: Models and methods*. Norwood, NJ: Ablex.

Orlady, L. M. (2000). *Automation issues: Perspectives from the flight deck* (2000-01-5596). Reston, VA: American Institute of Aeronautics and Astronautics.

Oser, R. (1999). A structured approach for scenario-based training. In *Proceedings of the Human Factors and Ergonomics Society 43rd annual meeting* (pp. 1138-1142). Houston, TX: HFES.

Oser, R., Cannon-Bowers, J., Salas, E., & Dwyer, D. (1999) Enhancing human performance in technology-rich environments: guidelines for scenario-based training. *Human/Technology Interaction in Complex Systems,* 9, 175-202.

Riggs, M.L. (1989). The development of self-efficacy and outcome scales for general applications. Paper presented at the *Society of industrial and Organizational Psychology Convention*, Boston, MA.

Scharr, T.M. (2002). Interactive video training for firearms safety. *Federal Probation, 65*(2), 45-51.

Shappell, S.A., & Wiegmann, D.A. (2001). Unraveling the Mystery of General Aviation Controlled Flight into Terrain Accidents Using HFACS. In *Proceedings of the 11th International Symposium on Aviation Psychology.* Columbus, OH: Ohio State University.

Sicard, B., Taillemite, J.P., Jouve, E., & Blin, O. (2003). Risk propensity in military and commercial pilots. *Aviation, Space, and Environmental Medicine, 74*(8), 879-881.

Simon, H.A. (1987). Decision making and problem solving. *Interfaces, 17*, 11-31.

Stewart, J.E., Dohme, J.A., & Nullmeyer, R.T. (2002). U.S. Army intial entry rotary-wing transfer of training research. *International Journal of Aviation Psychology, 12*(4), 359-375.

Stokes, A.F., Kemper, K.L., & Marsh, R. (1992). *Time-stressed flight decision making: A study of experts and novice aviators* (ARL-93-1/INEL-93-1). Arlington, VA: Office of Naval Research.

Sturgeon, W.R. (1988). *Ejection systems and the human factor: A guide for flight surgeons and aeromedical trainers* (88-TR-16). Downsview, Ontario: Department of National Defence – Canada, Defence and Civil Institute of Environmental Medicine.

Turner, T.P. (1995). *Cockpit resource management: The private pilot's guide.* USA: McGraw-Hill.

Weigmann, D., & Goh, J. (2000). *Visual Flight Rules (VFR) flight into adverse weather: An empirical investigation of factors affecting pilot decision making* (ARL-00-15/FAA-00-8). Savoy, IL: University of Illinois, Aviation Research Lab Institute of Aviation.

Wickens, C.D., Gordon, S.E., & Liu, Y. (1998). *An introduction to human factors engineering*. New York: Addison-Wesley Educational Publishers.

Wickens, C.D., Stokes, A, Barnett, B., & Davis, T., Jr. (1987). *A componential analysis of pilot decision making* (ARL-87-4/SCEEE-87-1). Savoy: University of Illinois, Aviation Research Laboratory.

Wilson-Donnelly, K.A., & Shappell, S.A. (2004). U.S. Navy/Marine Corps CRM training: Separating theory from reality. In *Proceedings of the Human Factors and Ergonomics Society* (pp.2070-2074). New Orleans, LA: HFES.

Zsambok, C.E. (1997). Naturalistic decision making: Where are we now? In C.E. Zsambok & G. Klein (Eds.), *Naturalistic Decision Making* (pp.3-16). Mahwah, NJ: Erlbaum.

Zsambok, C.E., & Klein, G. (Eds.). (1997). *Naturalistic Decision Making*. Mahwah, NJ: Erlbaum.

APPENDIXES A-F

Appendix A

Training Scenarios and Emergencies

Note to trainer: You will need to tell the pilot to explain aloud what he or she is thinking or observing throughout the experiment, in order to determine what he/she perceives and understands.
Encourage pilots to treat the simulation device like a real aircraft (e.g. react to an emergency in the same exact way, including using checklists)
For the first two scenarios, prompt pilots to use the 5 P's throughout different phases/decision points of the scenario. Then for the remaining 3 scenarios, tell pilot to continue to say their 5 P check but do not prompt them first each time. Take notes on how well the pilot utilizes the 5 P's during each scenario and give them feedback about their performance in the post flight discussions.
Note to technician: Set fuel selector on both for all scenarios. Hit "clear all weather" at the instructors station after every scenario.

Scenario 1 (20 minutes total)

Flight plan: Orlando International (KMCO) to Hernando County Airport (KBKV), FL (roughly 59 NM)
VFR conditions. Broken clouds from 5,000-9,000 ft. Northeast winds 8 knots. 4 pm in summer. Cruise at 3500 ft (MSL).
Realism setting off

Event 1 (10 minutes into flight)
Mild to moderate turbulence is encountered
At instructor's station: A few clouds at first, and then more cumulus clouds to overcast (5-9000 ft), with first mild then moderate turbulence under winds in advanced weather.

Performance measure:
Pilot should not lose control of the aircraft and continue to intended destination.

Pilot continues to Hernando County Airport.	Y N Wrong
Pilot maintains control of the aircraft.	Y N Wrong
If control is lost:	
Pilot has controlled emergency landing.	Y N Wrong
Pilot uses parachute.	Y N Wrong
If parachute is used:	
Pilot used parachute at correct time.	Y N Wrong
Pilot followed correct procedures (steps) for using parachute.	Y N Wrong
Pilot uses parachute at correct altitude and knots.	Y N Wrong
Pilot uses parachute over unpopulated area.	Y N Wrong
Pilot used the 5 P's in this event.	Y N Wrong

Event 2 (10 minutes into flight)
ATC alerts pilot that a level 4 thunderstorm cell is heading southeast towards the aircraft. As storm clouds gradually appear, severe turbulence is encountered.
At instructor's station: Add more clouds until reach thunderstorm conditions, then add severe wind (20 knots), then severe turbulence.

Performance measure:
Pilot should not lose control of aircraft and divert or turn away from storm.

Pilot diverts away from storm.	Y N Wrong
Pilot maintains control of the aircraft.	Y N Wrong
If control is lost:	
Pilot has controlled emergency landing.	Y N Wrong
Pilot uses parachute.	Y N Wrong
If parachute is used:	
Pilot used parachute at correct time.	Y N Wrong
Pilot followed correct procedures (steps) for using parachute.	Y N Wrong
Pilot uses parachute at correct altitude and knots.	Y N Wrong
Pilot uses parachute over unpopulated area.	Y N Wrong
Pilot used the 5 P's in this event.	Y N Wrong

Event 3 (15 minutes into flight/2 minutes after pilot diverts)
Due to pilots diverting/turning away from storm, a mid air collision occurs with another plane at 3500 ft AGL.
Note: Pilot should use BRS even if airport is nearby because the severe damage caused by a mid air collision would make it difficult to have a controlled landing.
At instructor's station: Tell pilot after diverting and about 2 miles from nearest airport. Hit 5 and 9 multiple times rapidly to simulate downward spin.

Performance measure:
Pilot should use parachute in less than 10 seconds from when they are told about the collision (about how long they would have before hitting the ground).

Pilot uses parachute.	Y N Wrong
Pilot used parachute at correct time.	Y N Wrong
Pilot followed correct procedures (steps) for using parachute.	Y N Wrong
Pilot uses parachute at correct altitude and knots.	Y N Wrong
Pilot uses parachute over unpopulated area.	Y N Wrong
Pilot used the 5 P's in this event.	Y N Wrong

Scenario 2 (25 minutes total)

Flight plan: Hanksville Airport (KHVE) in Hanksville, Utah to Canyonlands Field Airport (KCNY) in Moab, Utah (roughly 49 NM)
IFR conditions. Overcast from 7,000 to 12,000 ft. 1330 in winter. SW winds 2 knots. 15° F. Very light rain. Cruise at 10,000 ft MSL.
Realism setting ON
Reset instructor's station

Event 4 (15 minutes):
Flight continues and as pilot flies at about 3500 feet AGL, ice builds on the aircraft.
At instructor's station: Before this event – go to advanced weather – make sure clouds are overcast, add rain, then gradually change icing from moderate to severe until 10 minutes into scenario.

Performance measure:
Pilot should ask ATC to vector pilot out of ice. ATC will not be able to vector pilot out, so pilot must declare an emergency.

Pilot refers to checklist.	Y N Wrong
Pilot asks ATC to vector pilot out of ice.	Y N Wrong
Pilot declares an emergency.	Y N Wrong
Pilot used the 5 P's in this event.	Y N Wrong

Event 5 (20 minutes):
Heading indicator fails.
At instructor's station: Go to failures, heading indicator and hit failed then ok.

Performance measure:
Pilot should continue to the nearest airport or divert.

Pilot continues/diverts to nearest airport.	Y N Wrong
Pilot used the 5 P's in this event.	Y N Wrong

Event 6 (25 minutes)
Pilot should have unrecoverable stall from excessive ice. If pilot attempts to land, stalling should cause pilot to lose control on approach.
At instructor's station: Aircraft will continuously stall with too much ice.

Performance measure:
Pilot should probably use parachute (because will have difficulty making a controlled landing and will not be certain of the altitude of the terrain along the flight path, unless near airport).

Pilot does not allow to plane to drop below 6000 ft AGL.	Y N Wrong
Pilot has controlled landing.	Y N Wrong
If unable to land safely, pilot uses parachute.	Y N Wrong
If parachute is used:	
Pilot used parachute at correct time.	Y N Wrong
Pilot followed correct procedures for using parachute.	Y N Wrong
Pilot uses parachute at correct altitude and knots.	Y N Wrong
Pilot uses parachute over unpopulated area.	Y N Wrong
Pilot used the 5 P's in this event.	Y N Wrong

Event 7 (20-25 minutes/several miles from airport if pilot does have significant icing problems)
As pilot is flying in clouds, a complete electrical failure will occur.
At instructor's station: Fail electrical without restoration (SHIFT + T).

Performance measure:

Pilot refers to checklist.	Y N Wrong
Pilot has controlled landing.	Y N Wrong
If unable to land, pilot uses BRS parachute.	Y N Wrong
Pilot uses parachute at correct time.	Y N Wrong
Pilot uses parachute at the appropriate altitude and knots.	Y N Wrong
Pilot used the 5 P's in this event.	Y N Wrong

Scenario 3 (20 minutes total)

Flight path: Daytona International (KDAB) to Orlando International Airport (KMCO), FL (roughly 47 NM)
VFR conditions. 10 pm in summer. Cruise at 3500 ft (MSL).
In pre-brief: Carrying two passengers. One passenger is pregnant but is fit to fly.
Realism setting ON
Hit reset at instructor's station

Event 8 (15 minutes into flight/about 28 miles):
Intermittent engine interruption.

Performance Measure:
Pilot should continue and plan to land at the nearest airport (Sanford) while looking for places to land en route in case of complete engine failure.

Pilot refers to checklist.	Y N Wrong
Pilot continues to the nearest airport (Sanford).	Y N Wrong
Pilot does not use parachute.	Y N Wrong
Pilot used the 5 P's in this event.	Y N Wrong

Event 9 (20 minutes into flight/25.0 NM from destination):
Engine quits 5 miles before Sanford airport.
It should be too dark for pilot to find a place to find a place to land.
At instructor's station: Fail engine

Performance measure:
Pilot should glide plane to Sanford airport or nearest airport and make an emergency landing there, without using parachute unless absolutely necessary.

Pilot refers to checklist.	Y N Wrong
Pilot contacts ATC.	Y N Wrong
Pilot has controlled landing.	Y N Wrong
If unable to land safely, pilot uses parachute.	Y N Wrong
If parachute is used:	
Pilot used parachute at correct time.	Y N Wrong
Pilot followed correct procedures (steps) for using parachute.	Y N Wrong
Pilot uses parachute at correct altitude and knots.	Y N Wrong
Pilot uses parachute over unpopulated area.	Y N Wrong
Pilot used the 5 P's in this event.	Y N Wrong

Scenario 4 (10 minutes total)

Flight Plan: Centennial Airport (KAPA) in Denver to City of Colorado Springs Airport (KCOS), CO (roughly 49 NM)
VFR conditions. 9 pm in spring. Cruise about 9000 ft MSL
Realism setting ON

Event 10 (10 minutes into scenario):
Wing fire occurs in flight a little before 30 miles from Colorado Springs (10 miles before the Black Forest, so flying over heavily wooded area).
Note: Pilots will be told fire will not extinguish after refer to checklist and attempt resolve the situation. Make sure pilot knows to look for a road/airport first with clear conditions. If there is not one available, dive down to near 500 ft before using the parachute.
At instructor's station: Tell pilot there is a wing fire

Performance measure:
Pilot should refer to checklist and attempt to resolve the situation. Pilot should use parachute because will most likely not be able to land, as there are many trees below them.

Pilot refers to checklist.	Y N Wrong
Pilot has controlled landing.	Y N Wrong
If unable to land safely, pilot uses parachute.	Y N Wrong
If parachute is used:	
Pilot used parachute at correct time.	Y N Wrong

Pilot followed correct procedures for using parachute.	Y N Wrong
Pilot uses parachute at correct altitude and knots.	Y N Wrong
Pilot uses parachute over unpopulated area.	Y N Wrong
Pilot used the 5 P’s in this event.	Y N Wrong

Appendix B

Performance Assessment Scenarios: Pretest (30 min)

Note to raters: Please read over the following performance measures carefully before experiment. It is important to rate pilot consistently and their actions may not perfectly match one of the available choices on the 5 pt scale. For these reasons we ask for you to take notes if unable to reach a definite rating before the next event occurs, and then complete the rating after the test is finished.

**Go to Options – Flight Video – Record clip now

Scenario 1 (20 min total with preflight planning)

Flight plan: Kingston Ulster Airport (20N) in Kingston, NY to Stormville Airport (N69) in Stormville, NY (roughly 27 NM)
VFR conditions. 5 pm in summer. Cruise at 4500 ft MSL.

Event 1: (about 1 minute)
Immediately after takeoff, engine failure occurs (at only about 600 ft MSL)
At instructor's station: Go to failures, engine (or SHIFT + F)

Performance measure:
Pilot lands plane in a controlled manner if possible.

1- Pilot makes no attempts to land/Pilot loses control and crashed.
2-
3- Pilot makes unsuccessful attempt(s) to land.
4-
5- Pilot lands aircraft successfully.

If unable to land, pilot uses BRS parachute. **N/A**

1- Pilot does not use parachute.
2-
3- Pilot uses parachute incorrectly.
4-
5- Pilot uses parachute successfully.

Pilot uses parachute at correct time based on the sequence of events. **N/A**

1- Pilot deploys parachute excessively early or late.
2-
3- Pilot deploys parachute somewhat early or late.
4-
5- Pilot deploys parachute at the correct time.

Pilot uses parachute at the appropriate altitude. ______ ft N/A

1- Pilot uses parachute very inappropriately (well below 500 ft).
2-
3- Pilot uses parachute inappropriately (just below 500 ft).
4-
5- Pilot uses parachute very appropriately.

Pilot uses parachute at appropriate knots. ______ knots N/A

1- Pilot uses parachute very inappropriately (well over 90 knots).
2-
3- Pilot uses parachute inappropriately (just over 90 knots).
4-
5- Pilot uses parachute very appropriately (below 90 knots).

Overall, pilot performed the 5 P's/SRM behaviors.

1- Very inappropriately
2-
3- Inappropriately
4-
5- Very appropriately

Overall the pilot responded:

1- Very ineffectively.
2-
3- Ineffectively.
4-
5- Very effectively.

Pilot crashed. Y N

Performance Assessment Scenarios:
Posttest (about 2 hrs)

Note to raters: Please read over the following performance measures carefully before experiment. It is important to rate pilot consistently and their actions may not perfectly match one of the available choices on the 5 pt scale. For these reasons, we ask for you to take notes if unable to reach a definite rating before the next event occurs, and then complete the rating after the test is finished.
Note to technician: Tell participants to continue to recite 5 P's out loud.

Scenario 1 (30 minutes total with preflight planning)

Flight Plan: Daytona International (KDAB) to St. Augustine (KSGJ). (roughly 53 NM) VFR conditions. Broken clouds from 6-10,000 ft (add a brkn 6/7 cloud cell above scat 4/7) S SW winds 4 knots. 3 pm in summer. Cruise at 4000 ft.

Realism setting OFF

Event 1 (10 min into scenario):
After 10 min pass (or flying near Summer Haven), clouds gradually appear, signaling thunderstorm approaching. Severe turbulence is encountered. Pilot may attempt to fly around the clouds or turn around.
Nate: As ATC, contact pilot and let he/she know there is a level 4 T.S. cell ahead, moving in their direction (south). Approximate size of cell for pilot. Pilot may request permission to land at nearest airport, allow them to divert.
At instructor's station: Slowly add more cumulus clouds to broken over a couple of minutes (cell from 6,000 to 10,000 ft). Go to wind menu make wind speed 20 knots, with gusts of 25 knots. Also make turbulence severe. Change visibility to 10 miles for cloud cell and make thunderstorm conditions if they end up in storm. Even if pilot turns around add progressively worse winds and severe turbulence but have it gradually return to normal.

Performance measure:
Pilot continues to destination around storm or diverts to Flagler.

1- Pilot flies into storm.
2-
3- Pilot diverts/turns around after flying into storm.
4-
5- Pilot diverts/turns around immediately after stormy conditions appear.

Pilot maintains control of the aircraft.

1- Pilot loses control of the aircraft
2-
3- Pilot loses control of the aircraft, but then regains control.
4-
5- Pilot maintains control of the aircraft.

If pilot loses control of aircraft:
Pilot uses parachute. **N/A**

1- Pilot does not use parachute.
2-
3- Pilot uses parachute incorrectly.
4-
5- Pilot uses parachute.

Pilot uses parachute at correct time based on the sequence of events. **N/A**

1- Pilot deploys parachute excessively early or late.
2-
3- Pilot deploys parachute somewhat early or late.
4-
5- Pilot deploys parachute at the correct time.

Pilot uses parachute at the appropriate altitude. ______ ft **N/A**

1- Pilot uses parachute very inappropriately (well below 500 feet).
2-
3- Pilot uses parachute inappropriately (just below 500 feet).
4-
5- Pilot uses parachute very appropriately (above 500 feet).

Pilot uses parachute at appropriate NM. ______ knots **N/A**

1- Pilot uses parachute very inappropriately (well above 90 knots)
2-
3- Pilot uses parachute inappropriately (just above 90 knots).
4-
5- Pilot uses parachute very appropriately (below 90 knots).

Overall, pilot performed the 5 P's/SRM behaviors.

1- Very inappropriately
2-
3- Inappropriately
4-
5- Very appropriately

Overall, the pilot responded:

1- Very ineffectively
2-
3- Ineffectively
4-
5- Very effectively

Pilot crashed. Y N

Event 2 (12 minutes into scenario):
After flying around (or away from) the storm and a couple of minutes have passed, the airplane is level at 2,000 feet mean sea level (msl), at "cruise flight", then the airplane begins "pulling" to the left. The left wing has suffered damage due to the severe winds and turbulence. The pilot is over a populated area at this time.
At instructor's station: Add severe winds (200 knots) from 271°, then quickly switch back and forth between 0 and 100 knots. Also tell pilot: "You notice your airplane is pulling to the left, so you look out the window and see the turbulence has damage to your left wing (the current pulling is *not* caused by wind, but instead wing damage). The plane is marginally uncontrollable."

Performance measure:
Pilot uses parachute.

1- Pilot does not use parachute/lands plane.
2-
3- Pilot uses parachute incorrectly.
4-
5- Pilot uses parachute.

Pilot uses parachute at correct time based on the sequence of events. N/A

1- Pilot deploys parachute excessively early or late.
2-
3- Pilot deploys parachute somewhat early or late.
4-
5- Pilot deploys parachute at the correct time.

Pilot uses parachute at the appropriate altitude. ______ ft N/A

1- Pilot uses parachute very inappropriately (well below 500 feet).
2-
3- Pilot uses parachute inappropriately (just below 500 feet).
4-
5- Pilot uses parachute very appropriately (above 500 feet).

Pilot uses parachute at appropriate NM. ______ knots N/A

1- Pilot uses parachute very inappropriately (well over 90 knots).
2-
3- Pilot uses parachute inappropriately (just over 90 knots).
4-
5- Pilot uses parachute appropriately (under 90 knots).

Overall, pilot performed the 5 P's/SRM behaviors.

1- Very inappropriately
2-
3- Inappropriately
4-
5- Very appropriately

Overall, the pilot responded:

1- Very ineffectively.
2-
3- Ineffectively.
4-
5- Very effectively.

Pilot crashed. Y N

Scenario 2 (45 minutes total with preflight planning)

Flight Plan: Mt. Washington Regional Airport (KHIE) in Whitefield, NH to Oxford Municipal Airport in Oxford, ME (81B) (about 47 NM)
VFR conditions. Moderate north winds at 16 knots. Few clouds from 8,000-11,000 ft. 5:00 pm in winter (dusk). Cruise at 7500 ft MSL.
In pre-brief: Remind them that high winds are sometimes encountered near Mt. Washington. Explain northward winds but wind direction can vary. Let pilot know that

the FBO was closed before take off, so they did not refuel. Let them know they have the minimum fuel required (and not specifically how much).
Realism setting OFF

Event 3 (5 minutes into flight or when fly over Mt. Washington):
Extreme winds and tailwind as pilot passes Mt. Washington.
At instructor's station: Gradually add more wind from northward direction (360 degrees), increasing winds to 25 mph with 35 mph gusts and then reverse direction as go over mountain. Make turbulence occasional. Leave winds from southern direction. Gradually decrease conditions until just past mountain.

Performance measure:
Pilot maintains control of the aircraft.

1- Pilot loses control of the aircraft
2-
3- Pilot loses control of the aircraft, but then regains control.
4-
5- Pilot maintains control of the aircraft.

If pilot loses control of aircraft and is unable to regain it:
Pilot uses parachute. **N/A**

1- Pilot does not use parachute.
2-
3- Pilot uses parachute incorrectly.
4-
5- Pilot uses parachute.

Pilot uses parachute at correct time based on the sequence of events. **N/A**

1- Pilot deploys parachute excessively early or late.
2-
3- Pilot deploys parachute somewhat early or late.
4-
5- Pilot deploys parachute at the correct time.

Pilot uses parachute at the appropriate altitude. ______ ft **N/A**

1- Pilot uses parachute very inappropriately (well below 500 feet).
2-
3- Pilot uses parachute inappropriately (just below 500 feet).
4-
5- Pilot uses parachute very appropriately (above 500 feet).

Pilot uses parachute at appropriate NM. ______ knots N/A

1- Pilot uses parachute very inappropriately (well above 90 knots)
2-
3- Pilot uses parachute inappropriately (just above 90 knots).
4-
5- Pilot uses parachute very appropriately (below 90 knots).

Overall, pilot performed the 5 P's/SRM behaviors.

1- Very inappropriately
2-
3- Inappropriately
4-
5- Very appropriately

Overall, the pilot responded:

1- Very ineffectively
2-
3- Ineffectively
4-
5- Very effectively

Pilot crashed. Y N

Event 4 (25 minutes into flight/7 NM from destination or 4 NM if begin approach earlier)
Engine quits due to fuel leak.
Nate: If pilot contacts ATC at the closest airport and declares an emergency, clear them for immediate landing.
Instructor's station: Fail engine (SHIFT + F) permanently.

Performance measure:
Pilot refers to checklist to resolve problem.

1- Pilot does not refer to checklist.
2-
3- Pilot waits more than 2 minutes to refer to checklist.
4-
5- Pilot refers to checklist immediately.

Pilot follows checklist procedure to resolve problem.

1- Pilot does not follow checklist procedure.
2-
3- Pilot executes most steps correctly.
4-
5- Pilot executes all steps correctly.

Pilot contacts ATC.

1- Pilot does not contact ATC.
2-
3- Pilot waits until just before landing.
4-
5- Pilot contacts ATC immediately after using checklist.

Pilot declares an emergency.

1- Pilot does not declare an emergency.
2-
3- Pilot waits until just before landing.
4-
5- Pilot declares an emergency immediately after contacting ATC.

Pilot lands plane in a controlled manner if possible.

1- Pilot does not look for place to land before using parachute or crashing.
2-
3- Pilot makes unsuccessful attempt(s) to land.
4-
5- Pilot lands aircraft successfully.

If unable to land, pilot uses BRS parachute. **N/A**

1- Pilot does not use parachute.
2-
3- Pilot uses parachute incorrectly.
4-
5- Pilot uses parachute successfully.

Pilot uses parachute at correct time based on the sequence of events. **N/A**

1- Pilot deploys parachute excessively early or late.
2-
3- Pilot deploys parachute somewhat early or late.
4-
5- Pilot deploys parachute at the correct time.

Pilot uses parachute at the appropriate altitude. ______ ft N/A

1- Pilot uses parachute very inappropriately (well below 500 ft).
2-
3- Pilot uses parachute inappropriately (just below 500 ft).
4-
5- Pilot uses parachute very appropriately.

Pilot uses parachute at appropriate knots. ______ knots N/A

1- Pilot uses parachute very inappropriately (well over 90 knots).
2-
3- Pilot uses parachute inappropriately (just over 90 knots).
4-
5- Pilot uses parachute very appropriately (below 90 knots).

Overall, pilot performed the 5 P's/SRM behaviors.

1- Very inappropriately
2-
3- Inappropriately
4-
5- Very appropriately

Overall the pilot responded:

1- Very ineffectively.
2-
3- Ineffectively.
4-
5- Very effectively.

Pilot crashed. Y N

Scenario 3 (40 minutes total with preflight planning)

Flight plan: Four Corners Regional Airport (KFMN) in Farmington, NM to Cortez Municipal Airport (KCEZ) in Cortez, CO (about 38 miles)
IFR conditions. Overcast conditions from 6,500 to 12,000 ft MSL. Light rain. Some reports of icing in the forecast. Temperature is 0° F. 4 pm in winter. Visibility 3 miles. Cruise at 9500 ft MSL.
Realism setting ON
Reset instructor's station

Event 5 (15 minutes):
Ice gradually gathers on aircraft, stalling plane and degrading performance.
At instructor's station: Make icing severe immediately or if pitot heat is on wait longer, temp is 15° F and visibility is 3 miles. Drop base alt of clouds to 6000 ft 5 min into flight.
Nate: If pilot asks ATC to vector out of ice, say unable to (due to air traffic), so pilot must declare an emergency to get permission to divert/return to starting airport.

Performance measure:
Pilot refers to checklist to resolve problem.

1- Pilot does not refer to checklist.
2-
3- Pilot waits more than 2 minutes to refer to checklist.
4-
5- Pilot refers to checklist immediately.

Pilot follows checklist procedure to resolve problem.

1- Pilot does not follow checklist procedure.
2-
3- Pilot executes most steps correctly.
4-
5- Pilot executes all steps correctly.

Pilot asks ATC to vector pilot out of ice.

1- Pilot does not contact ATC.
2-
3- Pilot waits excessively long to ask ATC.
4-
5- Pilot contacts ATC immediately after onset of symptoms.

Pilot declares an emergency.

1- Pilot does not declare an emergency.
2-
3- Pilot waits too long to declare emergency.
4-
5- Pilot declares emergency immediately after ATC unable to vector pilot.

Overall, pilot performed the 5 P's/SRM behaviors.

1- Very inappropriately
2-
3- Inappropriately
4-
5- Very appropriately

Overall the pilot responded:

1- Very ineffectively.
2-
3- Ineffectively.
4-
5- Very effectively.

Pilot crashed. **Y** **N**

Event 6 (15+ min into scenario):
VOLTS light on annunciator flickers on and off, signaling alternator problems or problems with wiring.
Nate: If pilot contacts ATC, clear them to make an emergency landing at nearest airport.
At instructor's station: Press key for generator/alternator (SHIFT + A) several times for about a minute. Before this event make sure clouds are overcast, add rain, then gradually change icing from light to severe until 10 minutes into scenario.

Performance measure:
Pilot refers to checklist to resolve problem.

1- Pilot does not refer to checklist.
2-
3- Pilot waits more than 2 minutes to refer to checklist.
4-
5- Pilot refers to checklist immediately.

Pilot follows checklist procedure to resolve problem.

1- Pilot does not follow checklist procedure.
2-
3- Pilot executes most steps correctly.
4-
5- Pilot executes all steps correctly.

Pilot contacts ATC about temporary failures.

1- Pilot does not contact ATC.
2-
3- Pilot waits excessively long to contact ATC (e.g. a few minutes).
4-
5- Pilot contacts ATC immediately.

Pilot diverts flight to nearest airport.

1- Pilot makes emergency landing.
2-
3- Pilot continues on scheduled flight path.
4-
5- Pilot diverts to nearest airport.

Overall, pilot performed the 5 P's/SRM behaviors.

1- Very inappropriately
2-
3- Inappropriately
4-
5- Very appropriately

Overall the pilot responded:

1- Very ineffectively.
2-
3- Ineffectively.
4-
5- Very effectively.

Pilot crashed. Y N

Event 7 (20 minutes)
Pilot should have unrecoverable stall from excessive ice. If pilot attempts to land, stalling should cause pilot to lose control on approach.
At instructor's station: Aircraft will continuously stall with too much ice.

Performance measure:
Pilot lands plane in a controlled manner if possible.

1- Pilot makes no attempts to land/Pilot loses control and crashed.
2-
3- Pilot makes unsuccessful attempt(s) to land.
4-
5- Pilot lands aircraft successfully.

If unable to land, pilot uses BRS parachute. N/A

1- Pilot does not use parachute.
2-
3- Pilot uses parachute incorrectly.
4-
5- Pilot uses parachute successfully.

Pilot uses parachute at correct time based on the sequence of events. **N/A**

1- Pilot deploys parachute excessively early or late.
2-
3- Pilot deploys parachute somewhat early or late.
4-
5- Pilot deploys parachute at the correct time.

Pilot uses parachute at the appropriate altitude. ______ ft **N/A**

1- Pilot uses parachute very inappropriately (well below 500 ft).
2-
3- Pilot uses parachute inappropriately (just below 500 ft).
4-
5- Pilot uses parachute very appropriately.

Pilot uses parachute at appropriate knots. ______ knots **N/A**

1- Pilot uses parachute very inappropriately (well over 90 knots).
2-
3- Pilot uses parachute inappropriately (just over 90 knots).
4-
5- Pilot uses parachute very appropriately (below 90 knots).

Overall, pilot performed the 5 P's/SRM behaviors.

1- Very inappropriately
2-
3- Inappropriately
4-
5- Very appropriately

Overall the pilot responded:

1- Very ineffectively.
2-
3- Ineffectively.
4-
5- Very effectively.

Pilot crashed. Y N

Appendix C
BRS SRM KT

Participant # ________ Date ___/___/___

1. When should the BRS parachute be used?

2. When should the BRS parachute not be used?

3. What is the recommended minimum altitude for successful use of the BRS parachute?

4. What is the recommended maximum rate (in knots) for successful use of the BRS parachute?

5. You are flying a Cessna 172S from Daytona to St. Augustine in VFR conditions at an altitude of 3500 ft MSL and at a rate of 100 knots. It is 9 pm, and the engine cuts out. You are 7 miles from the airport, but you see some lights below. What would you do?

6. What are the six main behaviors of SRM?

    1.

    2.

    3.

    4.

    5.

    6.

7. Please circle the most appropriate response regarding whether or not each situation would most likely require *immediate* use of the BRS parachute:

Fluctuation in altimeter	Use BRS	Do not use BRS
Mid air collision	Use BRS	Do not use BRS
Structural failure	Use BRS	Do not use BRS
Engine failure near airport	Use BRS	Do not use BRS
Engine failure at night	Use BRS	Do not use BRS
Cabin fire	Use BRS	Do not use BRS
Stall on approach	Use BRS	Do not use BRS

8. What steps are involved in safe BRS activation?

1.

2.

3.

4.

5.

6.

7.

9. What are the 5 P's?

1.

2.

3.

4.

5.

10. What are the decision points for the 5 P's?

Appendix D
Self-efficacy Questionnaire

Participant # ________ **Date ___/___/___**

	Not at all true				Exactly true
1. I have confidence in my abilities to use the BRS parachute.	1	2	3	4	5
2. I believe I can become an expert at knowing when to use the BRS parachute in emergency situations.	1	2	3	4	5
3. I believe I can become an expert at performing single pilot resource management.	1	2	3	4	5
4. I know I can use SRM to help me through an emergency situation.	1	2	3	4	5
5. I feel I can make the best decision possible regarding whether to use the BRS parachute.	1	2	3	4	5
6. I have confidence in my abilities to save lives and the aircraft if at all possible in an emergency.	1	2	3	4	5
7. I feel I can make a good decision under pressure.	1	2	3	4	5
8. I am convinced that I will remember to use the BRS parachute before the minimum 500 ft altitude.	1	2	3	4	5
9. I am convinced that I will recall the correct procedure for using the BRS parachute.	1	2	3	4	5
10. I am confident in my ability to use the 5 P's in flight.	1	2	3	4	5

Appendix E

Participant # ________ **Date ___/___/___**

TLX

Place an X on the line in the position which best describes your evaluation.

1. How much mental and perceptual activity was required (e.g. thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?

   |--------|--------|--------|--------|--------|--------| N/A
   Low High

2. How much physical activity was required (e.g. pushing or pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

   |--------|--------|--------|--------|--------|--------| N/A
   Low High

3. How much time pressure did you feel due to the rate or pace at which the task or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

   |--------|--------|--------|--------|--------|--------| N/A
   Low High

4. How successful do you think you were in accomplishing the goals of the tasks set by the experimenter or yourself? How satisfied were you with your performance in accomplishing these goals?

   |--------|--------|--------|--------|--------|--------| N/A
   Low High

5. How hard did you have to work (mentally and physically) to accomplish your level of performance?

   |--------|--------|--------|--------|--------|--------| N/A
   Low High

6. How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed, and competent did you feel during the task?

   |--------|--------|--------|--------|--------|--------| N/A
   Low High

Appendix F
Demographics Questionnaire

Participant # ________ **Date ___/___/___**

______M ______F

Age: ______

Years of piloting experience: ______

Total flight hours: ______

Please list all aircraft licenses and ratings:

Have you ever flown an airplane or simulator equipped with a BRS parachute?

Have you ever declared an emergency when flying? If so, please explain.

Have you ever been involved in any aviation accidents or incidents as a pilot? If so, please explain.

Have you ever received training specifically for a Cirrus aircraft and/or the BRS parachute?

Please fill out this contact information. This information will be kept private, and will not be used in this experiment.

Name ____

Student Mailbox __________ Telephone Number _______________

Appendix G
Maximizing Pilot Performance

Conducted by Shayna Strally
Advisor: Elizabeth Blickensderfer
Embry Riddle Aeronautical University
Human Factors Research Laboratory
ERAU, Daytona Beach, FL 32114-3977

The purpose of this study is to examine aspects of pilot performance. The experiment consists of one session lasting approximately 5 hours. During this time you will fill out forms and receive training through a lecture format. You may also receive training from a flight instructor while flying an Elite™ flight simulation device. Tests in a flight simulator will be used to evaluate your performance.

Your performance scores will remain anonymous. There are no known risks associated with this experiment. You will be monetarily compensated for your participation, in the amount of $50. You may terminate your participation at any time. Your assistance will help us determine methods which may be used to maximize pilot performance.

Thank you for your participation. If you have any questions, please ask during the experiment or feel free to call me at (386) 226-4023, or Dr. Blickensderfer at (386) 323-8065.

Statement of Consent

I acknowledge that my participation on this experiment is entirely voluntary and that I am free to withdraw at any time. I have been informed as to the general scientific purposes of the experiment and that I will receive monetary compensation for completion of the study. If I withdraw from the experiment before its termination, I will not receive monetary compensation.

Participant’s name (please print): ____________________________

Signature of participant: ____________________________________ Date: ___________

Experimenter: ____ Date: ___________